J.J.C.Smart

Entre Ciencia y Filosofía

EF
ESTRUCTURA Y FUNCION
EDITORIAL TECNOS - MADRID

ESTRUCTURA Y FUNCION

El porvenir actual de la Ciencia

Colección dirigida por ENRIQUE TIERNO GALVÁN

1. QUENTIN GIBSON: *La lógica de la investigación social.*
2. M. F. ASHLEY MONTAGU: *La dirección del desarrollo humano* (2.ª ed.).
3. GUSTAV BERGMANN: *Filosofía de la ciencia.*
4. CHRISTIAN BAY: *La estructura de la libertad.*
5. ERNEST NAGEL: *La lógica sin metafísica.*
6. DAVID HILBERT y WILHELM ACKERMANN: *Elementos de lógica teórica* (2.ª ed.).
7. VERNON VAN DYKE: *Ciencia política: un análisis filosófico.*
8. KARL R. POPPER: *La lógica de la investigación científica* (2.ª ed.).
9. R. M. MARTIN: *Verdad y denotación.*
10. ERNEST GELLNER: *Palabras y cosas.*
11. C. O. BROAD: *El pensamiento científico.*
12. JOHN HOSPERS: *La conducta humana.*
13. ARTHUR PAP: *Teoría analítica del conocimiento.*
14. T. C. SCHELLING: *La estrategia del conflicto.*
15. RICHARD B. BRAITHWAITE: *La explicación científica.*
16. MILIC CAPEK: *El impacto filosófico de la física contemporánea.*
17. ROS ASHBY: *Proyecto para un cerebro.*
18. HANS REICHENBACH: *Moderna filosofía de la ciencia.*
19. GEORGE POLYA: *Matemáticas y razonamiento plausible.*
20. GEORGE L. S. SHACKLE: *Decisión, orden y tiempo.*
21. D. M. ARMSTRONG: *La percepción y el mundo físico.*
22. ERNEST NAGEL: *Razón soberana.*
23. MAX BLACK: *Modelos y metáforas.*
24. HASKELL B. CURRY y ROBERT FEYS: *Lógica combinatoria.*
25. LUDWIG WITTGENSTEIN: *Los cuadernos azul y marrón.*
26. ABRAHAM EDEL: *El método en la teoría ética.*
27. ALFRED LANDÉ: *Nuevos fundamentos de la mecánica cuántica.*
28. JEAN LADRIÈRE: *Limitaciones internas de los formalismos.*
29. N. WIENER y J. P. SCHADÉ: *Sobre modelos de los nervios, el cerebro y la memoria.*
30. G. H. VON WRIGHT: *Norma y acción.* Una investigacón lógica.
31. E. NAGEL y J. P. NEWMAN: *El teorema de Gödel.*
32. HENRY MARGENAU: *La naturaleza de la realidad física* Una filosofía de la Física moderna.
33. V. C. CHAPPELL: *El lenguaje común.* Ensayos de filosofía analítica.
34. ALF ROSS: *Lógica de las normas.*
35. W. SELLARS: *Ciencia, percepción y realidad.*
36. CHRISTIAN THIEL: *Sentido y referencia en la lógica de Gottlob Frege.*
37. CLAUDE FLAMENT: *Teoría de grafos y estructuras de grupo.*
38. W. y M. KNEALE: *El desarrollo de la lógica.*
39. G. E. HUGHES y M. J. CRESSWELL: *Introducción a la lógica modal.*
40. KARL R. POPPER: *Conocimiento objetivo.*
41. GOTTLOB FREGE: *Escritos lógico-semánticos.*
42. STEPHEN C. KLEENE: *Introducción a la Metamatemática.*
43. G. KALINOWSKI: *Lógica del discurso normativo.*

ENTRE CIENCIA Y FILOSOFIA

ESTRUCTURA Y FUNCION

EL PORVENIR ACTUAL DE LA CIENCIA

J. J. C. Smart
Universidad de Adelaida

ENTRE CIENCIA Y FILOSOFIA

Una introducción a la Filosofía de la Ciencia

EDITORIAL TECNOS

MADRID

Los derechos para la versión castellana de la obra
Between Science and Philosophy
publicada en inglés por
RANDOM HOUSE, Inc. Nueva York
© 1968 by Random House Inc., son propiedad de
EDITORIAL TECNOS, S. A.

Traducción por
MARIA ESPERANZA GUISAN SEIJAS

© EDITORIAL TECNOS, S. A., 1975
O'Donnell, 27 - Madrid-9
I.S.B.N. 84-309-0556-1
Depósito Legal: 3535 - 1957

Printed in Spain. Impreso en España por Selecciones Gráficas. - P.º de la Dirección, 52. - Madrid

En recuerdo de Janet

PREFACIO

Existen, por lo menos, tres tipos de textos o tratados sobre la filosofía de la ciencia. En primer lugar, está la filosofía de la ciencia del filósofo. Tales libros, probablemente estén llenos de rigor, quizá hagan uso de las técnicas de la lógica simbólica y, posiblemente, estén escritos, por lo menos en parte, para lectores competentes en ciencias, pero el argumento está a un nivel abstracto, tratando de ciertos conceptos elevados tales como la explicación, la ley de la naturaleza, y la confirmación, y no se interesan demasiado por el estado actual de las ciencias. En segundo lugar, está lo que podría denominarse filosofía de la ciencia "bebé": ésta intenta procurar un análisis filosófico de los actuales métodos científicos que resultará inteligible para los estudiantes de filosofía que no hayan recibido prácticamente formación científica alguna. En tercer lugar, está lo que podría denominarse filosofía de la ciencia del científico; ésta está más fundamentada en la ciencia actual que la filosofía de la ciencia del filósofo y presupone al menos una cierta preparación científica por parte del lector.

Este libro es el intento de una filosofía de la ciencia del científico no demasiado técnica, en el sentido de que está escrito para lectores que posean algo semejante al nivel de un estudiante de primer curso de matemáticas y física de los Estados Unidos o al nivel de los conocimientos de matemáticas del "A-level" de Inglaterra, lectores que puedan saborear revistas tales como la Scientific American. *No es propósito de esta obra competir con algunos de los excelentes tratados de la filosofía de la ciencia del filósofo (como los de Nagel, Hempel y Scheffler), pero intenta complementarlos. En los Estados Unidos, más que en Gran Bretaña o Australia, uno se encuentra con estudiantes de filosofía que tienen* por lo menos *este mínimo de preparación científica. Ello se debe a que los estudios de "letras" y "ciencias" no se realizan en facultades distintas de "letras" y "ciencias" como en Gran Bretaña y Australia. Sin embargo, incluso en estos últimos países uno se encuentra con* algunos *estudiantes que consiguen prescindir de la barrera entre letras y ciencias. (Varios de mis mejores alumnos no han venido a filosofía sino después de realizar un curso en ciencias o en ingeniería.) Este libro está escrito también para los científicos que se interesan por la filosofía. Pienso en particular en algunos de los científicos y tecnólogos que trabajan en, o cerca de, Adelaida, que han asistido a clases de filosofía en la universidad sólo por gusto. Probablemente debe haber muchos científicos como éstos interesados por la filosofía en todo el mundo.*

Un libro de filosofía se escribe inevitablemente desde un punto de vis-

ta particular. Ello se debe a que casi todas las cuestiones filosóficas son todavía objeto de controversia. Sin embargo, he intentado compensar los prejuicios personales incluyendo sugerencias respecto a lecturas posteriores al final de cada capítulo. Mis propios gustos personales cuentan también en la elección de las materias discutidas. El lector comprenderá, con toda seguridad, que existen muchos más problemas en la filosofía de la ciencia que los discutidos en este libro. Es de esperar, al menos, que las cuestiones discutidas aquí sean del interés del lector y que el libro le permita adentrarse en algunas de las más reñidas controversias contemporáneas y le lleve a su vez a interesarse por otras nuevas. El capítulo 2 está destinado a aquellos lectores con preparación científica pero que no han realizado estudios de lógica simbólica, como resumen de las ideas más importantes sobre la materia que puedan serle necesarias en los capítulos siguientes. Se trata nada más que de un resumen y no intenta sustituir a un tratado de lógica, de los que existen excelentes ejemplares. También he incluido algunas discusiones sobre la probabilidad en este capítulo.

Deseo dar las gracias a mi colega Mr. M. C. Bradley y a mi antiguo colega Mr. I. C. Hinckfuss, y también al profesor C. A. Hurst (del departamento de Física Matemática de la Universidad de Adelaida) por la lectura de un manuscrito previo de esta obra y por la aportación de valiosos comentarios. El profesor H. G. Andrewartha, del departamento de Zoología, Universidad de Adelaida; el profesor C. L. Hamblin, departamento de Filosofía, Universidad de Nueva Gales del Sur, y mi colega Míster S. E. Hughes, han tenido igualmente la gentileza de hacer comentarios sobre parte del libro. También quisiera dar las gracias a Mrs. Angela Bartesaghi, que mecanografió el manuscrito, y a Mr. Enrico Beretta, que me ayudó en el diseño de algunas de las gráficas.

J. J. C. SMART.
Universidad de Adelaida

INDICE

PREFACIO ... *Pág.* 9

Cap. 1. INTRODUCCIÓN ... 13

El ámbito de la Filosofía de la Ciencia ... 13
El ámbito de la Ciencia ... 15
La teoría de la estructura científica ... 16
La Ciencia y la Metafísica ... 18
La vindicación del método científico ... 21
Importancia práctica de la filosofía ... 22
Interacción entre los dos tipos de Filosofía de la Ciencia ... 23

Cap. 2. LÓGICA, SEMÁNTICA Y PROBABILIDAD ... 26

Lógica sentencial ... 26
Lógica de Predicados ... 28
Descripciones concretas ... 31
Teoría elemental de números ... 32
La teoría de Conjuntos ... 33
Semántica y verdad ... 35
Uso y mención ... 36
Contextos extensionales e intensionales ... 37
Probabilidad ... 38
La interpretación de la frecuencia de la probabilidad y la interpretación de la tendencia ... 42

Cap. 3. EXPLICACIÓN, LEYES Y TEORÍAS ... 47

Observaciones preliminares acerca de la explicación ... 47
¿Qué es la explicación? ... 49
El modelo de explicación deductivo-nomológica ... 50
Explicación estadística ... 55
Explicación incompleta ... 57
Esbozos de explicación ... 59
La explicación como analogía ... 59
La explicación como reducción a lo conocido ... 62
La explicación como redescripción ... 63
La crítica de Feyerabend al modelo deductivo ... 63
Crítica del uso, por parte de Feyerabend, del concepto de "Significado" (Meaning) ... 69
Los conceptos teóricamente neutrales y el lenguaje vulgar ... 70

Cap. 4. LA EXPLICACIÓN EN LAS CIENCIAS BIOLÓGICAS ... 74

Las ciencias históricas ... 74
Historia natural y taxonomía ... 75

La explicación en biología ... 78
Las explicaciones teleológicas y evolucionistas en la biología ... 80
Las Matemáticas en la Ecología y en la Teoría de la Evolución ... 85
Biología cósmica ... 86
Psicología ... 89

Cap. 5. Leyes y teorías en Física ... 96

Leyes experimentales, medida y aproximación ... 96
Las teorías consideradas como retratos (pictures) de la realidad ... 105
El instrumentalismo y el operacionismo ... 108
Crítica metafísica del instrumentalismo y el operacionismo ... 111
Crítica metodológica del instrumentalismo ... 120
La mecánica cuántica y el instrumentalismo ... 122
Leyes, teorías y lógica ... 126

Cap. 6. Inducción e hipótesis ... 135

El problema filosófico de la inducción ... 135
Objeciones de Duhem a la asimetría entre confirmación y refutación. 153
Conclusión ... 154

Cap. 7. Espacio y tiempo ... 158

Algunas teorías filosóficas acerca del espacio ... 158
La teoría del espacio y el tiempo de Newton ... 159
El ejemplo kantiano de la mano izquierda y la mano derecha ... 165
El espacio-tiempo en la relatividad especial ... 166
La paradoja del reloj ... 173
El espacio-tiempo en la relatividad general ... 180
Geometría y física ... 183

Cap. 8. La dirección del tiempo ... 194

El supuesto paso del tiempo ... 194
La asimetría temporal y la mecánica estadística ... 197
El principio de las ondas retardadas ... 209
La asimetría temporal y la expansión del universo ... 210

Cap. 9. Determinismo, libre albedrío e inteligencia ... 220

Determinismo e indeterminismo ... 220
El libre albedrío y el determinismo ... 225
Creatividad e inteligencia ... 230
Mecanismos intencionales ... 241
La psicolingüística y el problema kantiano ... 246

Notas referenciales ... 249

Indice alfabético ... 261

1. INTRODUCCION

El ámbito de la filosofía de la ciencia

Puesto que los filósofos difieren notablemente entre sí respecto al ámbito y la metodología de su disciplina, no es posible ofrecer una definición escueta y válida universalmente respecto al campo de la filosofía de la ciencia. Los desacuerdos acerca del ámbito de la filosofía llevan a desacuerdos acerca del ámbito de la filosofía de la ciencia. Pero, aun siendo posible una delimitación nítida del campo de la filosofía de la ciencia, no sería particularmente deseable. Supóngase que realizásemos tal delimitación y que un especialista en filosofía se encontrase con un problema de su interés, que se sintiese capaz de investigar, y que estuviese precisamente al otro lado de su campo así delimitado. Con toda seguridad que no desearíamos impedirle la investigación de este problema. Es importante impedir que la "departamentalización" demasiado rígida, que es característica de la estructura administrativa de las universidades, no departamentalice también con exceso nuestros objetivos y actividades intelectuales. Al igual que las diversas ciencias y la filosofía, la lógica y las matemáticas, progresan y se desarrollan, también cambian los modos en que es útil clasificar las diversas actividades intelectuales. Consideremos, por ejemplo, el atractivo campo de la moderna lingüística estructural y las investigaciones paralelas en la traducción mecánica de las lenguas. Un investigador de este campo se encontrará a sí mismo leyendo revistas dedicadas a filosofía, psicología, lógica matemática, teoría de la computabilidad, electrónica y filología. Tal reunión de disciplinas resultaría muy sorprendente tan sólo una generación atrás. Tales nuevas agrupaciones de las disciplinas proporcionan un terreno favorable para el filósofo, ya que al conjuntarse las materias seleccionadas de disciplinas anteriormente no relacionadas o apenas relacionadas, se originan problemas que él reconoce como típicamente filosóficos. En tal situación encontramos al filósofo aprendiendo algunas especialidades (tales como lingüística estructural o electrónica) y encontramos a los especialistas científicos interesándose por la filosofía.

Un modo de indicar el campo que en este libro se considerará "filosofía de la ciencia" sería decir que nos ocuparemos del tipo de cuestiones que se discuten en la revista americana *Philosophy of Science* o en *The British Journal for the Philosophy of Science*. Ciertamente si alguien que no fuese filósofo preguntase "¿Qué es la filosofía de la ciencia?", la forma más sencilla de contestarle sería decirle que fuese a la biblioteca y echase un vistazo a unos cuantos ejemplares de estas dos revistas. Si lo

hiciera se daría cuenta de que las cuestiones tratadas son de dos tipos principales, y, cuando no, son probablemente una mezcla de estos dos tipos. Por una parte encontramos discusiones *acerca de* la ciencia; entre ellas figuran las discusiones acerca de los modelos del argumento científico, los modos de probar las teorías científicas, la naturaleza de las leyes y teorías, y los modos de definir o introducir de alguna otra manera los conceptos científicos, lo que hace que una ley o teoría llegue a considerarse como más o menos "simple" que otra, y si el método científico en general puede ser defendido en contra de los argumentos escépticos que han venido apareciendo desde los tiempos de David Hume. Por otra parte, nos encontramos con discurso filosófico que no es tanto *acerca de* la ciencia, sino que *utiliza* la ciencia: puede ocurrir que los resultados de las investigaciones científicas nos ayuden a contestar algunas de las cuestiones acerca del hombre y del universo de las que se han ocupado tradicionalmente los filósofos. Por ejemplo, yo sostendría que los resultados científicos o bien son o bien serán relevantes respecto a la cuestión de si el espacio y el tiempo (o quizá el espacio-tiempo) debe considerarse como absoluto o relativo, y, asimismo, que las teorías científicas nos permiten comprender por qué el tiempo parece correr en una dirección. Y también, considerando otro campo de interés, las cuestiones acerca de la relación entre la mente y la materia y las del libre albedrío quizá resulten iluminadas por la moderna neurofisiología y por la cibernética (la teoría de la información y el control, incluyendo el estudio de los ordenadores).

En resumen, puede considerarse que la filosofía de la ciencia consta de dos componentes: 1) discurso analítico y metodológico *acerca de* la ciencia y 2) utilización de la ciencia como ayuda para la resolución de problemas considerados generalmente filosóficos. A la objeción de que el uso de la ciencia para solucionar un problema convierte a este problema automáticamente en científico, no filosófico, podemos reiterar nuestra oposición a la compartamentalización de las disciplinas académicas y nos podemos desinteresar de la cuestión de si a un problema se le denomina "científico" o "filosófico". No hay necesidad de un sindicalismo intelectual, y si un miembro de un departamento filosófico se siente a la vez interesado por una cuestión y capacitado para responderla no debe compungirse por no llevar bata blanca o mirar por un microscopio o por carecer del rótulo "departamento de física matemática" al final de su pasillo. Más aún, aunque una teoría científica pueda ser utilizada para tratar cuestiones que se consideran por lo común filosóficas, quizá no sea posible utilizarla de un modo directo ya que la cuestión posiblemente sea bastante especulativa o presente todo el tipo de peculiaridades y vaguedades intelectuales que por lo común plagan las discusiones filosóficas de cualquier tipo. Aunque se necesiten los resultados y los métodos científicos para tratar una cuestión, las habilidades científicas quizá no sean tan importantes como la clarividencia filosófica para deslindar un concepto de otro y para aplicar los resultados científicos. Esto no supone, por

supuesto, una rígida separación entre la filosofía y la ciencia, ya que los científicos, en ciertos campos por lo menos, tienen que llevar a cabo un gran número de innovaciones conceptuales y de aclaraciones antes de dedicarse a las matemáticas o a la química. La diferencia es sólo de énfasis. Por supuesto, por lo que a la innovación conceptual se refiere, podría muy bien argumentarse que los metafísicos han sido en gran medida sobrepasados en versatilidad y sutileza por los matemáticos y los físicos. Una razón más del por qué ciertas aplicaciones de la ciencia se consideran sin embargo filosóficas es la de que implican varias ramas distintas de la ciencia, y por lo tanto no son en modo alguno prerrogativa de ningún especialista particular.

El ámbito de la ciencia

El ámbito de la filosofía de la ciencia también depende del ámbito de la ciencia. Por ejemplo, algunos libros de filosofía de la ciencia se ocupan del análisis y la metodología de la explicación en general y, por tanto, su campo abarca no sólo el análisis conceptual y la metodología de las ciencias físicas y biológicas, sino también de la sociología y de la historia. En este libro excluiré la sociología y la historia y me limitaré a las ciencias físicas y biológicas. Gran parte de la historia es narrativa, y en cuanto va más allá de la narrativa y ofrece explicaciones, normalmente utiliza un conocimiento bastante vulgar acerca de los deseos, motivos y creencias humanos. Asimismo, la sociología se ocupa de hechos acerca de grupos particulares de personas, y al igual que la historia no suele utilizar como principios generales más que aquellos que son más bien de sentido común. Es cierto que complementa los principios del sentido común con otros menos evidentes que se derivan de análisis estadísticos de conjuntos de observaciones particulares, pero estos principios sólo pueden considerarse aproximadamente válidos para culturas y grupos particulares y distan mucho de ser leyes de la naturaleza. Por supuesto, muchos campos de estudio que por lo común se consideran ciencias, y que como tales serán considerados en este libro, son de tipo histórico en el sentido de que se ocupan de hechos particulares. Sin embargo, utilizan las ciencias físicas y biológicas de un modo en el que la historia y la sociología normalmente no las utilizan. Así, aunque la geología y la astronomía se ocupan de hechos particulares, utilizan teorías físicas y químicas muy complejas en su explicación de estos hechos particulares, y los hechos astronómicos se utilizan a menudo para comprobar teorías físicas. A su vez, la denominada teoría de la evolución es de tipo histórico, ya que se ocupa de un proceso terrestre particular (aunque de larga duración), pero está íntimamente ligada a las ciencias biológicas, tales como la genética y la bioquímica.

Por lo que a este libro se refiere, por tanto, consideraré que la "ciencia" comprende la física, la química, la astronomía, la geología y las ciencias biológicas, incluyendo la psicología experimental, pero no las

ciencias sociales y la historia. He estado dudando largo tiempo respecto a cómo clasificar la economía. He estado tentado a considerarla no como descriptiva, sino como una idealización comparable a la de la teoría matemática de los juegos, pero, por supuesto, en gran parte la economía pretende *ser* descriptiva y se ocupa de hechos históricos y su explicación. No parece contener leyes de la naturaleza del modo en que la física las contiene, y sin embargo se utiliza a los economistas a menudo como portadores profesionales de predicciones, de un modo en el que casi nunca se utiliza a los historiadores. Hasta la fecha parece faltar una filosofía de la economía que la sitúe correctamente y claramente en el mapa intelectual, por lo que queda una importante labor a realizar por parte de economistas con preocupaciones filosóficas y filósofos con preocupaciones económicas. En cualquier caso la economía no se considerará dentro de la esfera de este libro. Aunque no nos ocuparemos de la filosofía de la historia en cuanto tal, habrá, sin embargo, algunas discusiones acerca del modelo histórico de la explicación, ya que existen argumentos acerca de la evolución biológica, y en geología y astronomía que son, desde un punto de vista lógico, de carácter histórico, ya que proporcionan explicaciones de hechos particulares, aunque utilicen premisas generales derivadas de las ciencias y no dependan principalmente de generalizaciones de sentido común [1].

La teoría de la estructura científica

Existen algunos términos que son comunes en el discurso acerca de varias ramas diferentes de la ciencia, aunque no se dan esencialmente en el discurso acerca de la materia de estas ramas de la ciencia. Así, un científico que hable acerca de electrones, o genes, o estrellas quizá también necesite decir cosas acerca de sus enunciados *acerca de* los electrones, o los genes, o las estrellas. En este caso no está hablando acerca del tema de su ciencia, sino acerca de los enunciados de su ciencia. Por ejemplo, quizá necesite decir que éstos forman parte de una *teoría* o una *hipótesis,* que expresan una *ley,* o informan acerca de una *observación,* o que algunos de ellos son *evidencia* de otros. Considérense, pues, palabras tales como "teoría", "hipótesis", "ley", "enunciado de observación", "explicar", "evidencia". Estos términos aparecen en el discurso acerca de cualquier ciencia, al margen de la materia de que trate. Israel Schffler [2] * los ha denominado "términos estructurales". Se utilizan para describir la estructura del discurso científico. Aunque se dan dentro de las obras científicas, podríamos, sin embargo, imaginar que todas las explicaciones científicas tu-

[1] Desde luego, incluso la historia humana ha utilizado en algunas escasas ocasiones teorías científicas muy complicadas. Así, por ejemplo, el que una flota no consiguiese salir del puerto se ha explicado mediante una retroactividad astronómica de aguas muertas en la fecha en cuestión [1] *. Si este tipo de suceso fuese mucho más común, los historiadores habrían sido consideradores más científicos de lo que lo son ahora.

* Las notas de las referencias aparecen al final del libro, págs. 249-259.

viesen lugar sin ellos. Se dan únicamente cuando el científico se separa de su trabajo y nos dice lo que está haciendo, en lugar de continuar realmente haciéndolo. Es evidente, por ejemplo, que la tercera ley del movimiento de Newton podría haber sido establecida perfectamente aunque no hubiese sido jamás señalada con la palabra "ley" y la explicación genética de la relativa predominancia de anemia celular en los países afectados de paludismo podría haber sido ofrecida sin necesidad de haber sido designada con la palabra "explicación".

Obsérvese que estos términos son estructurales en el sentido en que se utilizan para *describir* la estructura, pero no para crear por sí mismos estructura. Es decir, no son estructurales exactamente en el mismo sentido en el que se les denomina a veces "estructurales" a los términos de la lógica, tales como "no", "o", "y", "si", "todos" y "algunos". Obsérvese también que aunque estos últimos términos de hecho aparecen por supuesto en todo discurso acerca de las distintas ciencias, también se dan dentro de las propias ciencias. Las expresiones lógicas son elementos estructurales *de* las sentencias, no términos descriptivos *acerca de* la estructura. Por este motivo quizá sea prudente el evitar el referirse a la investigación de la estructura científica, en el sentido de Scheffler, como "lógica de la ciencia", aunque se dé el caso de que muchos filósofos, especialmente algunos de aquellos influidos por Wittgenstein, quizá se sientan tentados a hacerlo (el propio Scheffler evita esta tendencia). Después de todo, en el sentido en que la palabra "lógica" será utilizada en este libro (y como se indicará en el capítulo 2), no existe ninguna lógica de la ciencia en particular. La ciencia utiliza la misma lógica que cualquier otro tipo de discurso, por ejemplo la matemática, la historia, la teología, la ética y la teoría política.

Schffler concibe una teoría de la estructura científica muy general, destinada a proporcionar la comprensión de los términos estructurales [3], y puesto que muchos de tales términos pertenecen no solamente al discurso acerca de las diversas ciencias sino también a parte del discurso ordinario (considérese la "hipótesis" o la "explicación" del detective) no existe una separación rígida entre la teoría de la estructura científica y la epistemología en general (teoría del conocimiento). Scheffler, por tanto, opera a un nivel de abstracción muy elevado. Siguiendo una sugerencia de Noam Chomsky, compara la relación entre, por una parte, la teoría de la estructura científica y, por otra parte, el uso de los términos estructurales en relación con las ciencias particulares, con la relación entre la teoría general de la lingüística estructural[2] y las gramáticas de las lenguas particulares. Por supuesto, al igual que las teorías de la lingüística estructural tienen que ser comprobadas mediante la consideración de su aplicabilidad a las diversas gramáticas de las lenguas particulares, del mismo modo la teoría de la estructura científica tiene que ser comprobada viendo cómo funciona en relación con el discurso más particular acerca de las expli-

[2] Campo al que Chomsky ha aportado importantes contribuciones [4].

caciones vulgares. Con todo, el hecho de que Scheffler opere a un considerable nivel de abstracción del discurso científico real aparece acentuado por el hecho de que los nombres de "Einsten", "Newton", "Darwin" y "Mendel" no figuren en el índice de su libro. Su libro no sufre menoscabo alguno por ello; de hecho, teniendo en cuenta sus objetivos y sus métodos filosóficos esta abstracción es un mérito. Sin embargo, una orientación más concreta también tiene su mérito, ya que existen muchos problemas que dependen de la apreciación de las *diferencias* entre las distintas ciencias y partes de la ciencia, y estas diferencias en la estructura pueden conducir a diferencias empíricas en los objetos que se discuten en estas ciencias, o partes de la ciencia. Así, como intentaré mostrar en el capítulo 4, existen diferencias estructurales y metodológicas entre la física y la biología que dependen de algunos hechos muy generales acerca del mundo. Incluso puede darse el caso de que las diferencias entre las explicaciones y las hipótesis de las distintas ciencias (o grupos de ciencias) presenten mayor importancia filosófica que las semejanzas. En cuyo caso una orientación más concreta, siguiendo más de cerca las teorías científicas reales, resultará valiosa.

La ciencia y la metafísica

Como observamos anteriormente, además de la filosofía de la ciencia concebida como análisis de la estructura y métodos de la ciencia, existe la filosofía de la ciencia concebida como aplicación del conocimiento científico a la solución de problemas filosóficos. Los capítulos 7-8 se centrarán principalmente en la filosofía de la ciencia de este segundo tipo. Si se admite de común acuerdo que la ciencia es lo que configura nuestra visión del mundo, debe admitirse, con toda seguridad, que la ciencia tiene algo que decirnos acerca de las grandes cuestiones relativas al hombre y al resto del universo de las que se han ocupado los metafísicos tradicionalmente. Por supuesto que a veces estas cuestiones pueden ser discutidas solamente de modo provisional y aproximado y que pueden depender de ramas de la ciencia bastante especulativas, como por ejemplo la cosmología. También pueden depender de algún tipo de síntesis de las ciencias especiales.

A partir de Ludwig Wittgenstein se ha creado un clima de opinión filosófica según el cual la ciencia es irrelevante para la resolución de los problemas filosóficos. Detengámonos, por tanto, a discutir el tipo de objeción que muchos filósofos presentarían en contra de las observaciones hechas en el parágrafo precedente. Será conveniente ampararnos en algunas agudas observaciones hechas por W. H. Watson en su libro *On Understanding Science* [5]. Watson es un físico que asistió a las explicaciones dadas por Wittgenstein en 1929-1931 y que estuvo muy influenciado por la filosofía de Wittgenstein. Su primera crítica de la filosofía sintética consiste en afirmar que si la filosofía fuese una especie de síntesis de las ciencias particulares entonces no decidiría nada: estaría basada en teorías

científicas particulares y serían los científicos los que decidirían si estas teorías debían ser aceptadas o si debían de ser descartadas como consecuencia de nuevos experimentos y observaciones [6]. La respuesta a esta crítica es con seguridad bastante simple, a saber, que el filósofo sintético no pretende "decidir las cosas" en este sentido. Su objetivo es el de relacionar entre sí las diversas ciencias, y su materia prima (las diversas teorías científicas existentes) la toma de los especialistas científicos. El no decide lo que sean estas teorías, pero sí pretende decidir cuál es el mejor esquema conceptual unificado mediante el cual pueden ser sintetizadas. Lo curioso del caso es que el filósofo sintético quizá se vea implicado en las mismas actividades que son recomendadas por aquellos filósofos "analíticos" hostiles a la idea de una filosofía sinóptica. Porque el filósofo sintético tendrá que aclarar los confusionismos conceptuales que se originan con frecuencia en las fronteras entre distintas disciplinas (donde ciencias distintas tienden a utilizar distintos tipos de conceptos para tratar de la misma materia). Es decir, el filósofo sintético tendrá que dedicarse a gran parte de la misma actividad general descrita y realizada por Gilbert Ryle en su libro *Dilemmas* [7], a saber, la investigación y resolución de disputas fronterizas. Por ejemplo, se argumentará en el capítulo 9, que es una confusión conceptual al relacionar distintas áreas del discurso lo que lleva a muchos filósofos a considerar que el principio de indeterminación de Heisenberg (uncertainty principle) tiene algo que ver con el problema del libre albedrío. Esta es la parte negativa de la filosofía sintética. El aspecto positivo, que los filósofos wittgensteinianos o analíticos niegan que sea parte de la filosofía, es la utilización de *otros* resultados científicos con relación al problema. Aunque podamos argumentar que algunos resultados científicos (el principio de indeterminación de Heisenberg, por ejemplo) *no* guardan relación con el problema del libre albedrío, también puede argumentarse que otros resultados científicos *sí* guardan relación con este problema.

Las consideraciones precedentes sugieren una nueva defensa de la idea de una filosofía sintética. Puesto que los especialistas de las diversas ciencias por lo común no suelen considerar las otras ciencias, parece haber un lugar en el mundo intelectual para el hombre que quiere, aunque sea de modo inexperto, observarlas todas y ver cómo se relacionan entre sí. Decir que tal hombre "no decide nada" no tiene más sentido que lo tendría decir que un físico teórico "no decide nada", puesto que es el experimentador el que proporciona los datos cruciales. Por supuesto, estoy dispuesto a admitir que el filósofo sintético es una persona de menor importancia que el físico. Ciertamente creo yo que los físicos teóricos han sobrepasado con mucho a los filósofos en su imaginatividad y en su capacidad de innovación conceptual y análisis. Pero admitir esto no implica negar que exista un lugar para el filósofo sintético: tratar de ver el mundo no como un físico o un químico, un biólogo o un psicólogo, sino como un todo completo, es con seguridad una tarea que tiene valor en sí misma.

No debemos pensar que lo único que necesitamos para lograr una síntesis filosófica es ser enciclopédicos: conjuntar simplemente los resúmenes de las distintas ciencias. De ser así, entonces no habría gran diferencia entre el filósofo sintético y el divulgador de la ciencia en general. Tal actividad meramente enciclopédica, sin embargo, no basta, debido a los problemas conceptuales que, como hemos indicado, se dan en la línea fronteriza entre distintas ciencias. Se necesita la habilidad de un filósofo para solucionar estos problemas. Podemos recordar la concepción que Ryle tiene del filósofo, como persona ocupada en la decisión de las disputas fronterizas entre ciencias distintas. Según Ryle, cuando dos campos aparecen en conflicto generalmente descubrimos que el conflicto no es real, y que cuando se han verificado las suficientes aclaraciones ambas partes pueden quedar satisfechas. En contra de esto yo opino que deberíamos inclinarnos más bien a suponer que uno de los contendientes puede simplemente resultar estar equivocado. Esto ocurre especialmente cuando el conflicto se da entre una ciencia y una creencia vulgar o tradicional. Por otra parte, las disputas fronterizas pueden darse incluso entre esquemas conceptuales dentro de una sola ciencia. Posteriormente discutiremos algunas de las dificultades que parecen originarse al relacionar la termodinámica fenomenológica y estadística [3]. Algunos científicos y filósofos desearían mantener que ambas formas de la termodinámica pueden ser válidas, y que los conflictos aparentes entre ellas se pueden resolver, mientras que otros dirían que la termodinámica estadística muestra, sencillamente, la falsedad de la termodinámica fenomenológica.

Sin embargo, Watson presenta una segunda crítica a la idea de una filosofía sintética que es más fuerte que su primera crítica, ya que de aceptarse ésta la primera resulta innecesaria. Dice Watson que cualquier enunciado o bien está "al mismo nivel" que un enunciado de un hecho científico o de lo contrario es un sinsentido. Aquí Watson está influido al parecer por el punto de vista expresado en el *Tractatus Logico-Philosophicus* de Wittgenstein [8]. En el *Tractatus* Wittgenstein mantenía que todas las sentencias significativas son imágenes (pictures) de hechos o complejos (funciones veritativas [4]) de imágenes de hechos. Esto conduce a la consideración de que la ciencia es una acumulación de imágenes, y que la síntesis de las ciencias consiste exactamente en una acumulación de acumulaciones de imágenes, es decir, exactamente en una gran acumulación. De acuerdo con la teoría de la imagen una innovación científica consistiría en la re-asignación de los valores de verdad ("verdadero" o "falso") que hemos asignado previamente a las sentencias, o quizá en asignar valores de verdad a sentencias que no habían sido consideradas anteriormente, o a las que no se les había asignado valores de verdad. Veremos, sin embargo, que el desarrollo de la ciencia no se ajusta a este modelo acumulativo o re-asignativo y que normalmente implica cambios radicales en los esquemas conceptuales. (Debe tenerse en cuenta que el

[3] Véanse págs. 63-68 más adelante.
[4] Véanse págs. 26-28 más adelante.

propio Wittgenstein llegó a descartar su punto de vista sobre el lenguaje mantenido en el *Tractatus,* y que su última posición no apoya la dicotomía "o bien enunciado de un hecho científico o de lo contrario un sinsentido".) El ensamblaje de los diversos esquemas conceptuales de la ciencia implica un separarse de los conceptos, un observarlos, o invirtiendo la metáfora del nivel de Watson, un ascenso a un nivel lingüístico más elevado donde momentáneamente tenemos que hablar acerca de las teorías de la física, la biología, etc., más que de los objetos de estas ciencias. Un buen ejemplo de un filósofo que hace un uso magistral de tal ascenso lingüístico temporal es W. V. Quine [9]. Este opina que un objetivo primordial de la filosofía es el descubrimiento del lenguaje que resulte más económico en sus compromisos ontológicos y que sea, sin embargo, capaz de expresar toda la ciencia. Lo cual se aproxima mucho al intento de hallar la teoría sinóptica más simple posible acerca del mundo.

Al insistir en el carácter atómico (o cuando menos discontinuo) de los hechos, Watson se opone a un escritor de la *Enclyclopaedia Britannica* [10], el cual, en el espíritu de la filosofía idealista, insiste en que el mundo no es una suma de hechos individuales "que existen yuxtapuestos unos a otros y que pueden ser conocidos aisladamente". De acuerdo con este escritor un hecho "no es nada sino en relación con otros hechos; y conforme estas relaciones se multiplican en el progreso del conocimiento, la naturaleza del denominado hecho se modifica indefinidamente". Despojando a esta afirmación de su terminología idealista, lo que viene a decir es que incluso las sentencias de observación están cargadas de teoría y no pueden comprenderse fuera de las concepciones teóricas con las que están cargadas. Si continuamos adelante esta línea de pensamiento no parece del todo absurdo que los hechos de las ciencias particulares puedan ser considerados de forma diferente cuando se contemplan desde el punto de vista de una sinopsis filosófica, o que la filosofía sintética pueda separarse de las ciencias particulares de las que parte, si bien de forma que puede ser de poca importancia *práctica* para los especialistas. La sugerencia no es tan absurda como pudiera parecer desde la perspectiva de una teoría de la "imagen" de cómo nuestro lenguaje representa al mundo. No hace falta que seamos idealistas para reconocer algunos buenos aciertos en los escritos de los filósofos idealistas. La crítica de Watson pierde su fuerza una vez que abandonamos la teoría de la imagen del lenguaje sobre la que parece estar basada.

La vindicación del método científico

En la sección precedente he intentado mostrar cómo puede defenderse el punto de vista de que la filosofía puede ser una especie de *superestructura* de las ciencias contra el tipo de objeciones presentadas por Watson. Examinemos ahora el punto de vista de que la filosofía también podría

proporcionar una especie de *subestructura* o fundamento de la ciencia. Este punto de vista es, con toda seguridad, absurdo, si con él se pretende que un físico o un geneticista, por ejemplo, deban sentirse inseguros hasta que hayan establecido un fundamento filosófico. Por supuesto, a menos que se desee que dividan sus energías, mejor será aconsejarles que ignoren el escepticismo filosófico. Sin embargo, esto no contradice el hecho de que un filósofo pueda estar interesado en preguntar por qué (como él probablemente cree) es la ciencia mejor que la magia o las brujerías. Por supuesto, resulta bastante evidente que existen buenas razones científicas para preferir la ciencia a la magia, pero también, es de suponer, habrá buenas razones mágicas para preferir la magia a la ciencia. ¿Sólo puede justificarse la ciencia frente a la magia de un modo patentemente circular? En el capítulo 6 nos encontraremos con dos modos posibles de justificar el método científico, "la validación" y "la vindicación". (Esta distinción entre validación y vindicación fue hecha por Herbert Feigl [11].) Se acepta generalmente que la validación del método científico es imposible, pero muchos filósofos mantienen que el otro tipo de justificación más débil, a saber la vindicación, sí puede lograrse. La cuestión de si esto es o no es verdad es un importante tema de estudio para el filósofo.

Importancia práctica de la filosofía

Nadie puede negar la importancia práctica de las ciencias, ya que son conocidas de todos tantas de sus interesantes aplicaciones tecnológicas. En comparación la filosofía (incluyendo la filosofía de la ciencia) parecería tener poca importancia práctica. Con todo sería desaconsejable pensar que la filosofía es un tema exclusivamente para habitantes de torres de marfil, y que por sí misma no posee valor práctico. La importancia práctica de la filosofía es triple: 1) En cuanto que es el arte de las aclaraciones conceptuales proporciona una habilidad para pensar claramente acerca de cuestiones poco claras. Esto será evidentemente de valor si tal habilidad puede transferirse al pensamiento ético, político y religioso, todo lo cual afecta, evidentemente, a los hombres de modo práctico. 2) En cuanto que algunas, cuando menos, de las aclaraciones conceptuales determinadas por el filósofo de la ciencia ayudan al científico a formular mejores teorías; este pensamiento filosófico tiene valor práctico indirecto en cuanto conduce al valor práctico de las teorías científicas en cuestión. 3) En cuanto que la filosofía es sinóptica y especulativa, puede tener efectos prácticos al sugerir las teorías científicas del futuro. Considérese el modo en que la teoría atómica metafísica de Demócrito debe haberse infiltrado en las culturas de épocas posteriores, facilitando así la invención de teorías científicas útiles y demostrables. Las especulaciones modernas acerca de las alternativas posibles a la interpretación de Copenhagen de la mecánica

cuántica [5] posiblemente pertenezcan a esta categoría de hipótesis indemostrables y todavía metafísicas que pueden servir, sin embargo, para producir teorías científicamente fructíferas. No obstante, no debemos olvidar el daño que puede hacérsele a la ciencia si nos dejamos fascinar en exceso por cualquier sistema especulativo filosófico, ya que puede impedir a nuestro pensamiento el deslindarse de un molde improductivo. Por ejemplo, antes de que los matemáticos consiguiesen que las geometrías no euclidianas fuesen tomadas en serio tuvieron que luchar contra prejuicios derivados de la filosofía kantiana. Albert Einstein, en sus "Autobiographical Notes" [12], dice que al inventar la teoría especial de la relatividad "El tipo de razonamiento crítico que se requería para el descubrimiento de este importante punto fue proporcionado decisivamente... especialmente por la lectura de los escritos filosóficos de David Hume y Ernest Mach". Podría argumentarse, sin embargo, que la influencia de Hume y March no fue del todo beneficiosa, puesto que los primeros escritos de Einstein sobre la relatividad están redactados de una forma innecesariamente operacionista [6] que es desorientadora e inconsistente con la posición filosófica posterior de Einstein, y por supuesto con la propia actividad científica de Einstein.

Interacción entre los dos tipos de filosofía de la ciencia

En este capítulo he distinguido dos tipos de filosofía de la ciencia: 1) análisis conceptual y metodológico de la ciencia, y 2), aplicación del conocimiento científico a la solución de problemas que han sido considerados filosóficos. Podría pensarse que existe escasa relación entre estos dos tipos de actividades a no ser la que se deriva del hecho de que ambas son realizadas por filósofos que se interesan por la ciencia y de que, por tanto, ambos tipos suelen aparecer en revistas tales como *Philosophy of Science*. Sin embargo, la relación es más íntima que ésta, ya que muchas de las actividades de la filosofía de la ciencia no son ni las del tipo 1) de forma pura, ni las del tipo 2) de forma pura, sino una mezcla de ambas. Por esta razón sería mejor describir 1) y 2) como dos *componentes* [7] de la filosofía de la ciencia. Es decir, con frecuencia tendremos 1) y 2) a la vez. Podemos tener 1) de forma pura pero no, en mi opinión, 2) de forma pura. Ello es debido a que un problema que se considera filosófico, y para el cual es relevante el conocimiento científico, difícilmente puede ser un problema directamente científico. Debe implicar elementos de confusión conceptual así como de problemática científica, ya que, de lo contrario, con toda probabilidad sería reconocido como cuestión científica, no como cuestión filosófica. Incluso aun cuando un problema fuese de

[5] Véanse págs. 122-126 más adelante.
[6] Sobre el oparicionismo véase el capítulo 5, más adelante.
[7] Análogos a los componentes de un vector.

gran interés emocional humano, como lo son los problemas del libre albedrío y la inmortalidad, con toda probabilidad no se consideraría filosófico si fuese evidente que pudiese resolverse dentro de una ciencia particular. Después de todo, las cuestiones formuladas por los biólogos acerca de las posibles causas del cáncer son por supuesto de gran interés humano, pero nadie las considera filosóficas. Lo característico es que se necesite aclaración conceptual antes de que el método puramente científico pueda ser aplicado a la solución del tipo de problema que consideramos filosófico. También puede darse una marcha atrás del conocimiento científico a la metodología. Por ejemplo en el capítulo 4 veremos que los hechos empíricos acerca de los organismos implican ciertas conclusiones acerca de la estructura conceptual y la metodología de las ciencias biológicas que las hace diferir en aspectos importantes de las ciencias físicas y de la química.

Sugerencias respecto a lecturas posteriores

En el texto se ha hecho referencia a Israel Scheffler, *The Anatomy of Inquiry* (New York: Alfred A. Knopf, 1963) y W. H. Watson, *On Understanding Physics* (London: Cambridge University Press, 1938). La obra de Scheffler es un valioso tratado relativo a la naturaleza de la explicación, los criterios de significado del lenguaje científico y la teoría de la confirmación. El libro de Watson es una aplicación original y difícil de algunas de las ideas de Wittgenstein a la aclaración conceptual de la física. Consecuencia de ello (o "remoto descendiente" como el autor dice) es la obra de W. H. Watson, *Understanding Physics Today* (London: Cambridge University Press, 1963). Excelentes tratados sobre la filosofía de la ciencia desde el punto de vista de la filosofía empirista y analítica son los de R. B. Braithwaite, *Scientific Explanation* (London: Cambridge University Press, 1953, y en rústica, New York: Harper and Row, 1960, traducción española: *La explicación científica,* ed. Tecnos, Madrid, 1964, y de Ernest Nagel, *The Structure of Science* (New York: Harcourt, Brace and World, 1961). Introducciones más elementales son las de C. G. Hempel, *Philosophy of Natural Science* (Englewood Cliffs, N. J.: Prentice-Hall, 1966); A. Pap, *An Introduction to the Philosophy of Science* (New York: Free Press, 1962), y Rudolf Carnap (Martin Gardner, ed.), *Philosophical Foundations of Physics, An Introduction to the Philosophy of Science* (New York: Basic Books, 1966). Una introducción elemental desde un punto de vista neo-wittgensteiano es la de Stephen Toulmin, *Philosophy of Science* (London: Hutchinson, 1953), y en rústica, New York: Harper and Row, 1960). Toulmin escribe para lectores no científicos y hace un ingenioso uso de ejemplos ilustrativos muy sencillos. Otro libro mucho más difícil, que está también escrito desde un punto de vista neo-wittgensteiano y que precisa más preparación científica por parte del lector, es el de N. R. Hanson, *Patterns of Discovery* (London: Cambridge University Press, 1958, y en rústica, New York: Cambridge University Press, 1965). La obra de K. R. Popper, *The Logic of the Scientific Discovery* (London: Hutchinson, 1959, y en rústica, ed. rev., New York: Harper and Row, 1965) es uno de los tratados más discutidos e influyentes escritos en el presente siglo (tradución española: *La lógica de la investigación científica,* ed. Tecnos, Madrid, 1962).

Unos cuantos útiles artículos sobre la filosofía de la ciencia han sido reunidos en la obra de H. Feigl y M. Brodbeck (eds.), *Readings in the Philosophy of Science*

(New York: Appleton-Century-Crofts, 1953) y en A. Danto y S. Morgenbesser (eds.), *Philosophy of Science* (New York: Meridian Books, 1960).

Ejemplos particularmente buenos de la filosofía de la ciencia en el sentido de aplicación de los resultados científicos a cuestiones filosóficas pueden verse en H. Reichenbach, *The Direction of Time* (Berkeley: University of California Press, 1956) y A. Grünbaum, *Philosophical Problems of Space and Time* (New York: Alfred A. Knopf, 1964).

Mi propia orientación filosófica ha variado considerablemente en los últimos diez años más o menos. En otro tiempo escribí un artículo, redactado en un espíritu muy semejante al punto de vista wittgensteiniano de W. H. Watson, y con el que ahora estoy en completo desacuerdo. Véase J. J. C. Smart, "The Relevance of Modern Analytic Philosophy for Science", *Australian Journal of Science,* XVI (1954), 165-17 y 215-218. Mi punto de vista actual puede verse en J. J. C. Smart, "Philosophy and Scientific Plausibility" en Paul K. Feyerabend y Grover Maxwell (eds.), *Mind, Matter and Method, Essays in Honor of Hebert Feigl* (Minneapolis: University of Minnesota Press, 1966), págs. 377-390. Una interesante crítica a la moderna orientación analítica y neo-wittgensteiana de la filosofía de la ciencia nos la ofrece P. K. Feyerabend en su polémico escrito: "Problems of Empericism" en R. G. Colodny (ed.), *Beyond the Edge of Certainty* (Englewood Cliffs, N. J.: Prentice-Hall, 1965), págs. 145-260. Wittgenstein no escribió mucho explícitamente sobre la filosofía de la ciencia, pero su primera posición es la de su *Tractatus Logico-Philosophicus,* y su última posición aparece mejor ejemplificada en su obra póstuma, *Philosophical Investigations* (Oxford: Blackwe, 1953). Antes de su muerte sus ideas sufrieron en gran parte la influencia de la filosofía británica y americana, a través de sus conferencias y de la influencia de sus alumnos. Un interesante artículo reciente sobre la relación entre la filosofía y la ciencia, que está escrito desde un punto de vista que es en algunos aspectos distinto de todos los considerados en el presente libro, es el de S. Körner, "Some Relations between Philosophical and Scientific Theories", *British Journal for the Philosophy of Science,* XVII (1966-67), 265-278.

2. LOGICA, SEMANTICA Y PROBABILIDAD[1]

LÓGICA SENTENCIAL

Considérense las sentencias "Jim es alto" y "Jim es pecoso". A partir de ellas podemos formar sentencias complejas tales como "Jim es alto o Jim es pecoso", "Jim es alto y Jim es pecoso" y "si Jim es alto entonces Jim es pecoso". También podemos utilizar "no" para formar una nueva sentencia a partir de otra sola, por ejemplo: "no (Jim es alto)" o de forma más coloquial "no es el caso de que Jim sea alto". En lugar de "no (Jim es alto)" en el lenguaje coloquial decimos "Jim no es alto", pero en lógica ponemos el "no" fuera de la sentencia. Es decir "no" en lógica debe leerse como "no es el caso de que". Debe observarse que el poner el "no" fuera de una sentencia puede implicar a veces el cambio total de la propia sentencia. Así, aunque "Jim no es alto" significa lo mismo que "no (Jim es alto)", la sentencia coloquial "algunos hombres no son altos" no puede ser traducida por "no (algunos hombres son altos)" (lo cual sería cierto si no hubiese hombres en absoluto) y debe ser traducida por la más complicada "hay algunos hombres y no (todos los hombres son altos)".

Las palabras "o", "y" y "si... entonces..." a las que añadimos "... si y sólo si..." son conectivas binarias de sentencias. "No" puede considerarse como una conectiva de sentencias unitaria. Estas conectivas de sentencias se escriben en símbolos lógicos como "$\cdot$" para "y", "$\vee$" para "o", "$\supset$" para "si... entonces...", "$\equiv$", para "... si y sólo si..." y "$\sim$" para "no". Pueden

TABLA 1

p	q	$\sim p$	$p \cdot q$	$p \vee q$	$p \supset q$	$p \equiv q$
V	V	F	V	V	V	V
F	V	V	F	V	V	F
V	F	F	F	V	F	F
F	F	V	F	F	V	V

[1] Gran parte de este capítulo consiste en un resumen de ciertas técnicas de la lógica, de la teoría de la prueba y de la semántica, que son necesarias para la comprensión del resto del libro. Va destinado en parte a los lectores científicos que carecen de estos conocimientos. Debe resaltarse que se trata simplemente de un resumen de conceptos y conclusiones relevantes para las partes del libro posteriores. Al lector interesado en detalles y pruebas se le recomienda que consulte tratados relevantes, algunos de los cuales aparecen citados en las Sugerencias para Lecturas Posteriores al final del capítulo. Los filósofos y los alumnos de Filosofía para quienes ésta es materia ya conocida puede pasarlo por alto. La parte final del capítulo acerca de la probabilidad es más debatible.

ser definidas mediante las siguientes tablas de verdad. Así la tabla de verdad para "$\supset$" proporciona los valores de verdad *(verdadero* o *falso)* de "$p \supset q$" para las diversas combinaciones de los valores de verdad de "p" y "q" (podemos considerar a "p", "q", etc., como sentencias mudas, y "$p \supset q$", por ejemplo, es un "esquema de sentencia" que se convierte en sentencia si "p" y "q" son sustituidas por sentencias adecuadas).

Así la tabla para "$p \supset q$" muestra que la sentencia compuesta es verdadera en todos los casos excepto cuando el componente de la izquierda es verdadero y el componente de la derecha es falso. Algunos filósofos dudan que "$\supset$" signifique lo mismo que "si... entonces..." en el lenguaje ordinario. Ciertamente podemos admitir que es extraño decir que "si 2+2=5 entonces existe un satélite de la Tierra". Sin embargo, lo extraño de esta expresión no muestra que posea un significado distinto del de "2+2=5 $\supset$ existe un satélite de la Tierra" (que es evidentemente cierto). La extrañeza puede derivarse simplemente del hecho de que carece de sentido afirmar una verdad tan evidente y de poca utilidad como "2+2=5 $\supset$ existe un satélite de la Tierra". En cualquier caso la cuestión de si las conectivas lógicas significan exactamente lo mismo que el lenguaje ordinario correspondiente carece de objetivo. Aquéllas realizan la mayoría de, y quizá todas, las funciones que un matemático o un científico necesitan que las conectivas realicen y están claramente definidas y exentas de ambigüedad.

Podemos construir esquemas de sentencias cada vez más complejos mediante repetidas aplicaciones de las conectivas de las sentencias. Así, considérese el esquema:

$$[\sim (p \cdot q) \vee (p \vee r)] \supset (q \supset \sim s)$$

Este se convierte en una sentencia determinada si las letras de la sentencia son sustituidas por sentencias. Algunos esquemas de sentencias deben tomar el valor de "verdaderos" con independencia de los valores de verdad de las sentencias que sustituyan a las letras de la sentencia. Así, ya bien sustituyamos la "p" de "$p \vee \sim p$" por una sentencia verdadera como "2+2=4", o por una falsa como "2+2=5", la combinación "$p \vee \sim p$" debe ser verdadera. (La verdad de "está lloviendo o no está lloviendo" no depende del tiempo atmosférico.) Los esquemas de sentencias que resultan verdaderos con independencia de los valores de verdad que asignemos a las letras de la sentencia se denominan esquemas de sentencias *tautológicos* o *válidos.*

Ejemplos más complejos de esquemas de sentencias válidos son:

1) $\sim (p \cdot q) \equiv (\sim p \vee \sim q)$
2) $[(p \supset q) \vee (p \supset r)] \supset [p \supset (q \vee r)]$

Existe un método mecánico (el método de las tablas de verdad) mediante el cual podemos determinar si un esquema de sentencia de la lógica proposicional es válido o no. Consideramos, simplemente, todas las posibles

combinaciones de los valores de verdad de las letras de las sentencias de nuestro esquema. Existen 4 combinaciones en el caso de 1) citado antes y 8 en el caso de 2). Si tenemos n letras de sentencia habrá 2^n combinaciones que tomar en consideración. Mediante reiteradas aplicaciones de las definiciones de la tabla de verdad de las conectivas de sentencias, para cada una de estas combinaciones, averiguamos el valor de verdad de la fórmula compleja para cada una de las combinaciones. Normalmente obtendremos una mezcla de Vs y Fs, pero si obtenemos V en todos los casos entonces el esquema es válido.

Lógica de predicados

Considérese las sentencias "Tom es alto", "Tom admira a María" y "Tom está entre Adelaida y Melbourne". Aquí "... es alto", "... admira a..." y "... está entre... y..." son, respectivamente, predicados monádicos, diádicos y triádicos. Un predicado, pues, es una expresión lingüística que contiene espacios en blanco de modo que produce una sentencia determinada cuando los espacios en blanco son sustituidos por nombres. Otra forma de convertir a un predicado en una sentencia determinada es utilizar los *cuantificadores* "algo" y "todo". Así, a partir de "... es alto" y "algo" obtenemos "algo es alto". En símbolos lógicos esto se escribe como "$(\exists x)$ x es alto". "Todo es alto" se escribe como "(x) x es alto". Entonces podemos expresar "Todo admira algo" como "$(x)(\exists y)$ x admira a y". "Algo admira todo" debe escribirse "$(\exists x)(y)$ x admira a y". "Algo se admira a sí mismo" debe ser "$(\exists x)$ x admira a x". "Todo está entre algo y algo" debe ser "$(x)(\exists y)(\exists z)$ x está entre y y z".

En "(x) x es alto" o en "$(\exists x)$ x es alto" se dice que la variable x está "ligada" mediante el cuantificador "(x)" o "$(\exists x)$". Una variable que no está ligada mediante ningún cuantificador (como "x" en "x es alto") se denomina "variable libre". Puede ser considerada como un nombre mudo. Por ejemplo, obsérvese que "$(\exists x)$ x es alto" y "$(x)(\exists y)$ x admira a y" difieren sólo notacionalmente de "$(\exists y)$ y es alto" y "$(z)(\exists w)$ z admira a w" —o también de "$(y)(\exists x)$ y admira a x"—. Por otra parte "x es alto" e "y es alto" son no sólo distintos notacionalmente simplemente. No podemos sustituir siempre impunemente "x" por "y". Por ejemplo "x es alto $\supset$ x es alto" da lugar evidentemente a una sentencia verdadera sustituyamos x por lo que lo sustituyamos (y por tanto es un esquema válido), pero "x es alto $\supset$ y es alto" no es un esquema válido ya que da lugar a una falsedad si sustituimos "x" por el nombre de algo que es alto e "y" por el nombre de algo que es corto. Debe observarse que se precisa algún método de colocación de paréntesis, evidentemente, para indicar el "ámbito" (scope) de un cuantificador. "$(x)(Fx \supset Gx)$" tiene evidentemente un significado distinto del de "$[(x)Fx] \supset Gx$" en el cual la última aparición de x es libre. El último esquema difiere sólo notacionalmente de "$[(y)Fy] \supset Gx$", mientras que "$(x)(Fx \supset Gx)$" no es en modo alguno equivalente a "$(y)(Fy \supset Gx)$".

Ahora sustituyamos los predicados "... es alto", "... admira...", "... está entre... y..." por letras predicativas "*F*", "*G*", "*H*", etc. De este modo podemos formar expresiones tales como "*Fx*", "*Fxy*", "*Fxyz*". Mediante reiteradas aplicaciones de cuantificadores y letras predicativas podemos formar una expresión tal como, por ejemplo, "$(\exists x)(y)[(Fxy \supset Gz) \cdot Hxyz] \supset (\exists w) Gw$". (Esta contiene una variable libre "*z*" y las demás variables son variables ligadas.) Podemos realizar una *interpretación* de este esquema sustituyendo las variables libres por nombres y las letras predicativas por predicados apropiados. De este modo "*G*" podría ser interpretada como "... es alto", y "*F*" como "... admira a...". Ahora supongamos, por ejemplo, que todos los pares de cosas del universo ordenadas de tal modo que la primera admira a la segunda son exactamente los mismos pares ordenados de tal modo que la primera ve a la segunda. Si se diese tal caso, los predicados "... admira..." y "... ve..." serían extensionalmente equivalentes: serían "verdaderos con respecto a" los mismos pares ordenados de cosas. En lógica de predicados consideramos que los predicados extensionalmente equivalentes son el mismo predicado, y podemos especificar un predicado simplemente especificando con respecto a qué cosas del universo es verdadero (o respecto a qué pares ordenados de cosas, triples ordenados de cosas, etc., en el caso de predicados diádicos, triádicos, etc.). Obsérvese que mientras que las sentencias son "verdaderas" o "falsas", los predicados son "verdaderos respecto a" o "falsos respecto a". La clase de cosas respecto a la cual un predicado es verdadero se denomina a veces la "denotación" o "extensión" de aquel predicado.

Podemos definir ahora la validez en la lógica de predicados. Se dice que una fórmula de la lógica de predicados es válida si resulta verdadera en todo universo no vacío para todas las interpretaciones de sus letras predicativas. Por ejemplo, la fórmula "$(x)(\exists y) Fxy \supset (\exists x)(y) Fxy$" no es válida. Considérese un universo consistente en individuos *a* y *b*, tal que "F" sea verdadero respecto a los pares ordenados (*a*, *b*) y (*b*, *b*) y falso respecto a los pares ordenados (*b*, *a*) y (*a*, *a*). (Para captar esto intuitivamente obsérvese que "alguien ama a todo el mundo" no se sigue de "todo el mundo ama a alguien".) Por otra parte, no existe ningún universo ni ninguna interpretación de letras predicativas que haga que

$$\text{"}(x)(Fx \supset Gx) \supset (y)[(\exists x)(Fx \cdot Hyx) \supset (\exists x)(Gx \cdot Hyx)]\text{"}$$

resulte falso [2].

Es posible especificar ciertos axiomas y reglas a partir de los cuales se pueden derivar *todos* los esquemas válidos de la lógica de predicados. Tales conjuntos de axiomas y reglas se dice que están *completos*. La completitud (completeness) de un conjunto de axiomas y reglas de la lógica de

[2] Una sentencia que presenta la forma de este último esquema es la de "Si todos los canguros son marsupiales entonces todas las cabezas de canguros son cabezas de marsupiales". No hace falta que sepamos zoología o anatomía alguna para asegurarnos de la verdad de esta sentencia: la sola lógica garantiza su verdad.

predicados fue demostrada por Kurt Gödel en 1930 [1]. Cuando se ha demostrado que un sistema de axiomas y reglas está completo normalmente es fácil mostrar lo mismo respecto a otros sistemas determinados demostrando que el primer sistema puede ser derivado de los últimos.

Por otra parte la lógica de predicados se diferencia notablemente de la lógica sentencial en que no existe procedimiento o método "efectivo" mediante el cual podamos determinar si un esquema arbitrario es válido o no. Si tenemos suerte (o inteligencia) podemos dar con una prueba de la validez o la no validez pero no existe ningún algoritmo que nos asegure el éxito. (Compárese la diferencia entre la invención de una prueba geométrica y la realización de una división larga. Lo primero requiere suerte e inteligencia mientras que lo segundo requiere la simple aplicación de un mecanismo.) Es cierto que existe un método mecánico para la construcción de una secuencia de fórmulas válidas de la lógica de predicados, de forma que si aplicamos el método durante bastante tiempo aparecerán finalmente todas las fórmulas válidas. Por otra parte, no existe ningún método mecánico semejante para construir una secuencia mediante la cual aparecerán finalmente todas las fórmulas no válidas. Por tanto, existe un método mecánico que nos proporcionará finalmente la respuesta "Sí" a la pregunta "¿Es ésta una fórmula válida?", con tal que la respuesta *sea* "Sí". Por otra parte, si una fórmula no ha aparecido todavía no sabemos si el no haber sido capaces de construirla se debe al hecho de que no hemos aplicado nuestro método durante el tiempo suficiente o si por mucho que apliquemos nuestro método nunca construiremos nuestra fórmula. Podríamos contestar a nuestra pregunta de forma mecánica sólo si existiese un método de construir, una tras otra, todas las fórmulas no válidas. Sin embargo, es imposible matemáticamente que pueda existir tal método, al igual que es matemáticamente imposible cuadrar el círculo. En otras palabras, la lógica de predicados se diferencia de la lógica sentencial en que no existe procedimiento de decisión. No sólo no conocemos ningún procedimiento de decisión tal, sino que no existe ningún método de decisión en absoluto [3].

La imposibilidad de un procedimiento de decisión en la lógica de predicados fue mostrada por Alonzo Church en 1936 [2]. Esto tiene consecuencias importantes para las matemáticas y la ciencia. La lógica de predicados (a diferencia, por ejemplo, de la lógica sentencial, a la cual contiene como una parte propia) contiene en sí misma todos los recursos necesarios para expresar *cualquier* argumento matemático o científico. Es decir, cualquier teoría científica completamente axiomatizada puede ser expresada mediante las notaciones de la lógica de predicados, siempre que las letras predicativas sean sustituidas por predicados constantes adecuados. Sim-

---

[3] Ciertos *fragmentos* de la lógica de predicados sí cuentan con procedimientos de decisión. Así la lógica de predicados monádicos (que contiene únicamente fórmulas con predicados monádicos) cuenta con un procedimiento de decisión. También cuenta con él el conjunto de sentencias de la lógica de predicados en las que los cuantificadores universales proceden a todos los existenciales.

plemente añadimos los axiomas, en la forma expresada, de la ciencia en cuestión a los axiomas y reglas de la propia lógica de predicados. Tomando un ejemplo muy sencillo, la teoría de la identidad puede ser expresada mediante la adición de los axiomas:

1) $x=x$
2) $(x=y) \supset (y=x)$
3) $(Fx \cdot x=y) \supset Fy$

Aquí "=" es un predicado diádico constante que significa "es idéntico a". Ahora bien, puesto que cualquier teoría matemática o científica, axiomatizable, finita, puede ser expresada dentro de la lógica de predicados en la forma acabada de citar, un procedimiento de decisión de la lógica de predicados llevaría a un modo mecánico de decidir si alguna sentencia de la teoría era o no era un teorema. Para la validez habría que comprobar simplemente la fórmula $A \supset T$, donde A es la conjunción de los axiomas y T es el teorema putativo. Todas las teorías matemáticas o científicas parecen ser axiomatizables (si bien todavía no se ha llevado a cabo el proceso de axiomatización de la mayor parte de las teorías y en algunos casos sólo podría ser realizado tras arduo trabajo), y por tanto una prueba de un procedimiento de decisión para la lógica de predicados haría trivial, en cierto sentido [4], toda la solución de problemas en matemáticas. El teorema de Church muestra que nunca se dará esta trivialización.

Se da el caso, sin embargo, de que algunos fragmentos de la lógica y la matemática sí admiten un procedimiento de decisión. Por ejemplo Tarski [3] demostró que el álgebra de los números reales cuenta con un procedimiento de decisión [5]. (En el álgebra de los números reales las variables son nombres mudos que ocupan el lugar de los números reales y no existen cuantificadores. Y no hace falta que mencionemos las *clases* de los números reales como en el análisis.)

Descripciones concretas

La lógica de predicados con la identidad nos permite actuar sin nombres o expresiones de una sola referencia. Es decir, podemos sustituir una sentencia tal como "el satélite natural de la Tierra es esférico" por una sentencia que no contenga ninguna expresión tal como "el satélite natural de la Tierra". En su lugar escribimos:

$$(\exists x)\,[x \text{ es esférico y } (y)\,\{y \text{ es un satélite natural de la Tierra} \equiv (x=y)\}]$$

Una ligera reflexión mostrará que la sentencia anterior establece que una

[4] Seguirían siendo necesarios métodos de solución de problemas *en la práctica* si se diese un procedimiento de decisión que fuese demasiado largo y complicado para ser utilizado en la vida real.

[5] El procedimiento no es lo suficientemente rápido para ser utilizado en computadores reales. Mi amigo C. L. Hamblin está trabajando en la adaptación del método de Tarski a usos prácticos.

y una sola cosa es un satélite natural de la Tierra y que esta cosa es esférica. En general escribimos:

$$F\,[(\imath x)\ Gx] = \text{def.}\ (\exists x)\,[(y)\ Gy \equiv (x=y)\cdot Fx]$$

Una frase de la forma "$(\imath x)\ Gx$" se denomina descripción concreta, y puede leerse como "la G-adora". (Si capta o no capta exactamente el significado de frases de la forma "el tal y cual" del lenguaje coloquial no interesa, pero cuando menos se aproxima a ello.) Un mecanismo debido a W. V. Quine nos permite sustituir siempre nombres por descripciones concretas. Leemos simplemente "Sócrates" como "el socratizador", donde socratizar equivale a poseer el conjunto de propiedades que nos permite diferenciar a Sócrates. "El socratizador" suena artificial, pero en principio no difiere de la sustitución de la "Luna" por "el satélite natural de la Tierra".

Teoría elemental de números

La teoría elemental de números es la teoría de los enteros, 0, 1, 2, 3... No comprende fracciones y números reales, ni tampoco clases de enteros. La teoría elemental de números puede ser formalizada dentro de la lógica de predicados mediante la adición de axiomas apropiados. Ahora bien, la sintaxis de una teoría formalizada, que trate con nociones tales como "fórmula bien formada" y "deducible de" (y por tanto con "demostrable", ya que esto significa "deducible de los axiomas"), puede ser reproducida en la teoría elemental de números. Es decir, existe un modo de asignar números naturales a los símbolos del sistema, y a las secuencias de símbolos del sistema, de modo que se dan diversas relaciones aritméticas entre estos números si y sólo si se dan relaciones sintácticas correspondientes entre las expresiones correspondientes a estos números. El método de correlacionar números con expresiones, secuencias de expresiones y secuencias de secuencias de expresiones, se denomina numeración de Gödel. La teoría elemental de números es la teoría matemática más débil en la que puede reproducirse la sintaxis. Obsérvese que la sintaxis trata sólo de símbolos y de los modos de manipularlos, y que no trata en absoluto de su interpretación. Aprovechándose de esto, y reproduciendo la sintaxis de la teoría de números en la propia teoría de números, Gödel demostró cómo construir una sentencia G, de un sistema formalizado de la teoría de números, que es la sentencia aritmética que corresponde a la afirmación sintáctica de que la sentencia en cuestión es indemostrable en el sistema, si el sistema es consistente. Por supuesto, si el sistema es inconsistente cualquier fórmula puede ser demostrable, así como su negación. Ya que las reglas de la lógica sentencial nos permiten pasar de "p" y "$\sim p$" a "q", y por tanto mediante sustitución a una proposición cualquiera. (Prueba: a partir de "p" obtenemos "$p \vee q$", que

junto con "$\sim p$" da lugar a "q".) También presuponiendo la consistencia [6] puede mostrarse que la negación de G tampoco es demostrable. También puede mostrarse (en un argumento dentro de la *semántica* del sistema en cuestión, véase pág. 35) que aunque G es indemostrable es, sin embargo, verdadera.

Por supuesto se pueden aumentar los axiomas de la teoría de números de forma que G resulte demostrable. (Es trivial que esto podría hacerse añadiendo G como axioma.) Pero entonces podría mostrarse con referencia al conjunto de axiomas aumentado que si éste es consistente entonces habría una sentencia de la teoría elemental de números G' que no sería ni demostrable ni indemostrable. Y así sucesivamente. Los axiomas de la teoría elemental de números son incompletables, al menos si la teoría elemental de números es consistente. Más aún, cualquier teoría formalizada más potente, tal como la teoría de conjuntos, que sea lo suficientemente poderosa para que pueda derivarse de ella la teoría de números, debe ser también, si es consistente, incompletable. Puesto que la sintaxis puede ser reproducida en la teoría de números, *a fortiori* puede ser reproducida en esta teoría más potente, y así el argumento de Gödel puede ser desarrollado con respecto a ella.

La teoría de conjuntos

La teoría de conjuntos puede ser axiomatizada dentro de la lógica de predicados mediante la incorporación del predicado diádico constante "ϵ" ("es un miembro de") y axiomas apropiados. El predicado "es idéntico a" y la noción de equivalencia entre conjuntos puede definirse rápidamente. Dos conjuntos son equivalentes si sus elementos están entre sí en relación biunívoca. (A pesar de la etimología la noción de relación "biunívoca" puede ser definida *antes* de que sea definido el número natural.) Por ejemplo, el conjunto de números naturales y el conjunto de números pares son equivalentes, ya que existe una relación biunívoca que empareja x y $2x$. Las propias relaciones son conjuntos, ya que son definidas como pares ordenados, y los propios pares ordenados se identifican con los conjuntos. De este modo el par ordenado (x, y) puede ser identificado con $\{\{x\}, \{xy\}\}$, es decir, el conjunto que contiene el conjunto cuyo único miembro es x y el conjunto cuyos miembros son x e y.

Resulta ser que toda la matemática, incluyendo el análisis, puede ser reproducida en la teoría de conjuntos, o dicho más ampliamente, "puede ser expresada" en la teoría de conjuntos. Por ejemplo, en una versión de la teoría de conjuntos (la de Zermelo) los números naturales se identifican con Λ, $\{\Lambda\}$, $\{\{\Lambda\}\}$, etc., donde Λ es el conjunto vacío. En otra (la de Von Neumann) los números son Λ, $\{\Lambda\}$, $\{\Lambda, \{\Lambda\}\}$, etc., de forma que el

[6] En realidad Gödel presupuso un requisito más fuerte que la consistencia, denominado consistencia-ω. Rosser comprobó una versión del teorema de Gödel utilizando solamente el requisito más leve de la mera consistencia.

número n es el conjunto de todos los números menos n. Existen todavía otras versiones. Carece evidentemente de significado preguntar si el número 2 "es realmente" $\{\{\Lambda\}\}$ o $\{\Lambda\ \{\Lambda\}\}$, o también cualquier otra entidad. Cualquier secuencia de objetos de la teoría de conjuntos que formen una progresión realizarán la tarea por lo que al matemático se refiere.

Debe recordarse que aunque los números naturales *no* se identifiquen con entidades de la teoría de conjuntos, la teoría de conjuntos tiene un importante papel que desempeñar en matemáticas, ya que el análisis requiere la consideración de conjuntos infinitos de números reales.

Si todos los miembros de un conjunto A son miembros de un conjunto B, entonces se dice que A es un subconjunto de B. Obsérvese que un conjunto es subconjunto de sí mismo. Es fácil mostrar que el conjunto de subconjuntos de un conjunto dado (ya sea finito o infinito) no puede ponerse en correspondencia biunívoca con aquel conjunto. Esto es, por supuesto, muy evidente en el caso de conjuntos finitos: por ejemplo, el conjunto de subconjuntos del conjunto de dos miembros $\{a, b\}$ tiene cuatro miembros: Λ, $\{a\}$, $\{b\}$, $\{a, b\}$. Si dos conjuntos son tales que uno es equivalente a un subconjunto de otro, pero no a la inversa, se dice que el primero es menor que el segundo. Así el conjunto de los números naturales es menor que el conjunto de subconjuntos del conjunto de números naturales. Este último conjunto es menor que el conjunto de subconjuntos de sí mismo. El conjunto de los números naturales tiene un número $\aleph_0$. Puede demostrarse fácilmente que el conjunto de los subconjuntos del conjunto de los números naturales es equivalente al conjunto de los números reales. Así pues, el número de este conjunto, $2^{\aleph_0}$, es el número del continuo. ¿Es $2^{\aleph_0}$ el menor número transfinito después de $\aleph_0$? Es decir, ¿existe un número transfinito, distinto de $\aleph_0$, que sea menor que $2^{\aleph_0}$? La hipótesis del *continuo* es que no existe: es decir, que el número transfinito del continuo es el primero después de $\aleph_0$. Gödel demostró que la hipótesis del continuo es consistente con los otros axiomas de la teoría de conjuntos [4]. Más recientemente Paul J. Cohen [5] demostró que su negación es consistente también. Esto proporciona otro ejemplo del hecho de que la cuestión relativa a qué teoría de conjuntos adoptamos es en gran medida cuestión de gustos: parece ser que hay teorías de conjuntos alternativas e igualmente buenas en un sentido en que no hay lógicas alternativas e igualmente aceptables, o, por lo que a esta cuestión se refiere, hipótesis físicas alternativas e igualmente aceptables.

Algunas de estas materias distan mucho de ser problemas de la filosofía de la ciencia, pero en la pág. 186 necesitaremos cuando menos la distinción entre $\aleph_0$ y $2^{\aleph_0}$. La teoría de conjuntos siempre presentará *cierto* interés para los filósofos de la ciencia, ya que la física precisa de la teoría de las variables reales (análisis) y esto sólo puede discutirse con la ayuda del discurso acerca de los conjuntos infinitos de los números reales.

Finalmente, el enunciado de que la teoría de conjuntos puede ser axiomatizada mediante la adición de ciertos axiomas y el predicado constante "ϵ" a la lógica de predicados, requiere una ligera aclaración. En algunas

versiones de la teoría de conjuntos se necesitan reglas adicionales de formación (reglas de tipos en Russell; reglas de estratificación en Quine). Estas reglas evitan que construyamos expresiones que de otro modo nos conducirían a inconsistencias. Considérese la paradoja de Russell "el conjunto de todos los conjuntos que no son miembros de sí mismos es un miembro de sí mismo". Si se nos permite suponer que existe tal conjunto entonces resulta una paradoja, ya que si es miembro de sí mismo no lo es y si no lo es lo es. Reglas de tipos o de estratificación nos impiden construir la sentencia transgresora, ya que se convierte en "no gramatical". En otros sistemas los propios axiomas aseguran la no existencia del conjunto transgresor, sin necesidad de reglas de formación adicionales.

Semántica y verdad

Se ha mencionado que la sintaxis de una teoría puede ser reproducida en esa misma teoría, siempre que la teoría sea, cuando menos, tan potente como la teoría elemental de números. Sin embargo, esto no es verdad con relación a la semántica que trata no sólo de nociones tales como "fórmula bien formada", "demostrable" y "consistente", sino también de nociones tales como "denota", "es verdadero respecto a" y "verdadero". Que una sentencia sea verdadera depende, en general, tanto del modo en que esté el mundo como de las reglas del lenguaje. Así "la nieve es negra" sería verdadera siempre que "negra" se usase con la interpretación que damos en el presente a "blanca". Así no debemos decir que una serie de símbolos es verdadera *simpliciter,* si no que es verdadera en algún lenguaje *L.* "La nieve es blanca" es verdadera en español si y sólo si la nieve es blanca. Existe la tentación de generalizar esto diciendo que en general "'*p*' es verdadera" si y sólo si *p.* Sin embargo, esto es incorrecto, ya que "*p*" no es un nombre variable de una sentencia sino que es el nombre de la letra "*p*". A fin de vencer esta dificultad Alfred Tarski [6] ha definido, en efecto, "verdadero" en términos de "verdadero respecto a". Un predicado "*F*" es verdadero respecto a *x* si y sólo si *Fx.* Entonces mediante la consideración del modo en que una sentencia es construida en un lenguaje formalizado dado a partir de sus predicados, junto con los cuantificadores y las conectivas de sentencias, Tarski fue capaz de definir "verdadero en *L*", siendo *L* un lenguaje formalizado. Debe observarse que puesto que el aparato formal de la teoría tiene que ser considerado específicamente en la construcción de la definición de verdad, esto proporciona una razón matemática de por qué lo que se ha de definir no es "verdadero" sino "verdadero en *L*". Tarski también demostró que "verdadero en *L*" puede definirse únicamente en un lenguaje con medios de expresión más potentes que los de *L.* En particular, se necesita la teoría de conjuntos para definir "verdadero en *L*" cuando *L* es un sistema formalizado de la teoría elemental de números. Al usar "verdadero" y "falso" como predicados de *sentencias,* estamos presuponiendo que las sentencias

se utilizan independientemente del contexto. En el lenguaje coloquial, desde luego, a menudo utilizamos sentencias dependientes del contexto. Así, si una persona dice "soy alto" quizá diga algo verdadero, mientras que otra persona puede utilizar la mismísima sentencia para decir algo falso. En estos casos quizá deseemos hablar de la verdad y la falsedad de las *expresiones verbales* o *escritas* de las sentencias más que de las propias sentencias. Sin embargo, para los objetivos de este libro presupondremos que las sentencias de una teoría científica se expresan con independencia del contexto.

La teoría de la prueba requiere, en general, tanto de la sintaxis como de la semántica. Así, que un teorema sea demostrable o no a partir de los axiomas de un sistema formal es cuestión de si ciertas series de símbolos pueden ser transformadas en otras de acuerdo con ciertas reglas. Se pueden considerar los símbolos simplemente como si fuesen piezas de ajedrez que hubiesen de ser movidas de acuerdo con las reglas de un juego. Hasta aquí nos movemos dentro del campo de la sintaxis. Sin embargo, cuando preguntamos si todas las fórmulas *verdaderas* o *válidas* son demostrables estamos utilizando, evidentemente, conceptos de la semántica.

Uso y mención

La sintaxis y la semántica, debe observarse, son disciplinas en las que hablamos *acerca de* el simbolismo de un sistema. La distinción entre *usar* un símbolo o serie de símbolos y *mencionar* este símbolo o serie de símbolos no debe ser olvidada nunca por los filósofos y lógicos. Por ejemplo, en la lógica sentencial podríamos utilizar "*p*", "*q*", "*r*"... como letras de sentencias y utilizar "$A \supset B$" como el *nombre* mudo (o variable) de una serie de símbolos que consta de: 1), alguna expresión construida a partir de "*p*", "*q*", "*r*"..., y las conectivas "$\sim$", "$\cdot$" "$\supset$", etc., seguida de 2), el símbolo "$\supset$", seguido de 3), alguna expresión construida a partir de "*p*", "*q*", "*r*"..., y las conectivas de sentencias. En filosofía también es importante tener continuamente en cuenta la diferencia entre usar una expresión lingüística y mencionarla. Por ejemplo, en la filosofía de las matemáticas no se debe olvidar la distinción entre *numerales* y *números*. Los primeros son expresiones lingüísticas mientras que los últimos son entidades matemáticas no lingüísticas. Los expertos en lógica quizá harán la observación de que de haber sido estrictamente riguroso yo debería haber usado letras metalingüísticas apropiadas (tales como "*A*" y "*B*", anteriormente en este parágrafo) en vez de "*p*" y "*q*" en la tabla de verdad de la pág. 26 más atrás, ya que la tabla estaba destinada a mostrar cómo funcionan las conectivas de sentencias entre cualquier combinación bien formada de letras de sentencias y conectivas de sentencias y no simplemente entre las propias letras "*p*" y "*q*". Sin embargo, allí he adoptado el

método menos riguroso (pero común) de utilizar "p" y "q", sacrificando así el rigor en aras de la simplicidad de la exposición.

El error común de interpretar "$\supset$" como "implica" tiene lugar por la confusión entre uso y mención [7]. "$\supset$" es una conectiva de sentencias: une dos sentencias para formar una sentencia compuesta y pertenece al mismo nivel que "y" y "o". "Implica" no es una conectiva sino un predicado binario, que une *nombres* de sentencias. Obsérvese que "Sócrates es hombre implica Sócrates es mortal" no es ni siquiera buena expresión gramatical, ya que ninguna sentencia contando con sólo estas palabras puede contener tres verbos. Debemos decir "'Sócrates es hombre' implica 'Sócrates es mortal'" o "la proposición *de que* Sócrates es hombre implica la proposición *de que* Sócrates es mortal" [7].

Lo que significa ordinariamente "implica" es oscuro: posiblemente sería mejor construir "'—' implica '...'" diciendo que la sentencia "..." es deducible (quizá junto con alguna teoría o contando con algún conjunto de informaciones de "—", o quizá debe construirse como afirmación de la *validez* de "—$\supset$...". Pero "'—' implica materialmente '...'" está perfectamente claro: afirma simplemente la verdad de "—$\supset$...". Es un enunciado de la semántica acerca de las sentencias "—" y "...", a saber, que cuando éstas van unidas por "$\supset$", apareciendo la primera de ellas primero, dan lugar a una sentencia verdadera. La proposición "(x) (x está loco $\supset$ x tiene alguna deficiencia bioquímica en sus neuronas)" es verdadera o falsa según la extensión del predicado "está loco" esté incluido o no en la extensión del predicado "tiene alguna deficiencia bioquímica en sus neuronas". Puede observarse que la cuestión de si términos semánticos como "verdadero" y "denota" se aplican correctamente o no en un caso particular puede dar lugar a cuestiones sobre hechos empíricos acerca del mundo.

Contextos extensionales e intensionales

Considérese la expresión "no se da el caso de que" o "$\sim$". Si sustituimos A en $\sim A$ por alguna otra sentencia B cuyo valor de verdad sea el mismo que el de A (o, cuando A es un esquema de sentencia que contiene variables libres, si sustituimos A por algún esquema de sentencia B que sea verdadero con relación a los mismos objetos acerca de los que A es verdadero) entonces el valor de verdad de $\sim A$ es el mismo que el de $\sim B$. Podemos denominar a "$\sim$" "contexto extensional" de la expresión sobre la que opera. Ahora considérese el contexto "es una cuestión de pura matemática el que...". Es una cuestión de pura matemática el que $2+2=4$, pero no es cuestión de pura matemática el que la Tierra tenga

[7] Decimos a veces que dos sentencias expresan la misma proposición, de modo que podemos considerar que una proposición es aquello con lo que todas estas sentencias mantienen determinada relación. Algunos filósofos, como por ejemplo Quine, consideran imprecisa la noción de la expresión de una misma proposición y prefieren hablar únicamente de sentencias e inscripciones o expresión verbal de las sentencias.

exactamente un satélite natural, aun cuando las sentencias "2+2=4" y "la tierra tiene exactamente un satélite natural" tengan el mismo valor veritativo. Decimos que "es una cuestión de pura matemática el que" es un operador, o conectiva de sentencias, intensional, no extensional.

Nuevamente sea *A* una sentencia que contenga uno o más nombres o expresiones de una única referencia. Sean sustituidos estos nombres, o expresiones de referencia única, por otros que señalen los mismos objetos, de modo que produzcan una sentencia *B*. Entonces en un lenguaje extensional *A* será verdadera si y sólo si *B* lo es. Por ejemplo, "no se da el caso de que el número de los planetas visibles sea menor que diez" es verdadera si y sólo si no se da el caso de que cinco sea menor que diez (ya que el número de los planetas visibles es de hecho cinco). Sin embargo "es una cuestión de pura matemática el que el número de los planetas visibles sea menor que diez" es falsa, mientras que "es una cuestión de pura matemática que cinco es menor que diez" es verdadera. Así pues, "no se da el caso de que..." es un contexto extensional y "es una cuestión de pura matemática el que" es intensional o modal.

He utilizado "es una cuestión de pura matemática el que..." en mi ejemplo a fin de evitar expresiones como "lógicamente necesario" o "lógicamente posible", o incluso "necesario" o "posible" solamente, debido a la existencia de controversias filosóficas relacionadas con estas palabras. Sin embargo "necesariamente" y "posiblemente" son, desde luego, contextos intensionales típicos. A veces el lenguaje intensional puede ser sustituido por lenguaje extensional en el metalenguaje: por ejemplo, el *operador* intensional "necesariamente" en "necesariamente 2+2=4" puede ser sustituido por el *predicado* extensional "es necesario" aplicado a la sentencia "2+2=4", como en "'2+2=4' es necesario". Se originan problemas técnicos y filosóficos cuando se combinan operadores intensionales con la lógica cuantificacional. Afortunadamente si una teoría científica puede ser formalizada puede ser formalizada en un lenguaje extensional. (Trataré de defender este punto de vista con cierta amplitud en este libro, pero en general el lector debe considerar esto como un presupuesto en el que está basado este libro.)

Probabilidad

Algunos filósofos, al parecer, consideran a la teoría de la probabilidad como una rama de la lógica que estudia una relación que es más débil que la de la implicación (o quizás la deducibilidad). En este caso cometen el error de confundir la lógica con su metateoría (sintaxis o semántica). Los símbolos en los que la teoría de la probabilidad se expresa a veces pueden ser los que originen esta confusión. Supóngase que, tal como a veces ocurre, una fórmula típica de la teoría de la probabilidad se escribe más o menos como sigue:

$$P(p \cdot q, r) = P(p, r) \times P(q, p \cdot r)$$

¿Cómo podríamos leer esta fórmula? Supóngase que la leemos diciendo que la probabilidad de que tanto *p* como *q* sean verdaderos, dado que sea *r* verdadero, es igual a la probabilidad de *p* dado *r* multiplicado por la probabilidad de *q*, dados *p* y *r*. Obsérvese que esto tiene sentido sólo si consideramos a "*p*", "*q*" y "*r*" como *nombres* mudos de sentencias y no como las propias sentencias, como en la lógica sentencial. En otras palabras, la teoría de la probabilidad debe ser interpretada no como análoga a la lógica, sino como análoga a la metalógica.

Para evitar confusiones, por tanto, sustituyamos "*p*", "*q*", "*r*", etc., por las letras "*a*", "*b*", "*c*", etc., y quede entendido que una posible interpretación de la fórmula

$$P(a \cdot b, c) = P(a, c) \times P(b, a \cdot c)$$

es metalingüística, en la que *a*, *b*, *c* son *nombres* mudos de sentencias. Y evitaremos la confusión que se originaría al denominar a esta interpretación "lógica", denominándola interpretación "*metalingüística*".

En ciencias, sin embargo, la palabra "probabilidad" se utiliza para hablar no acerca de sentencias sino acerca de sucesos del mundo. Así la fórmula precedente sería interpretada como: "La probabilidad de que un suceso que es del tipo *c* sea un suceso que es a la vez del tipo *a* y del tipo *b* es igual a la probabilidad de que un suceso *c* sea un suceso *a* multiplicado por la probabilidad de que un suceso *b* sea a la vez un suceso *a* y un suceso *c*".

Sea *c* la clase de estudiantes de una universidad. Sea *a* la clase de los estudiantes de ciencias y sea *b* la clase de los estudiantes del sexo femenino. Entonces la fórmula $P(a \cdot b, c) = P(a, c) \times P(b, a \cdot c)$ puede interpretarse como diciendo que la proporción de estudiantes universitarios que son a la vez del sexo femenino y de ciencias es igual a la proporción de estudiantes que estudian ciencias multiplicado por la proporción de estudiantes de ciencias que son del sexo femenino.

Sería posible, en el caso de la universidad, examinar a todos los estudiantes y ver cuántos eran del sexo masculino, cuántos del femenino, cuántos de ciencias y cuántos no de ciencias. En la ciencia normalmente nos ocupamos de conjuntos de sucesos, o de cosas, tales que de ellos sólo puede ser observada una muestra. Considérese el hecho de arrojar una peseta al aire. Puede hacerse esto un número indefinido de veces, y si se consigue una larga secuencia se verá que la proporción de pesetas que salen "cara" se acerca a $\frac{1}{2}$ (si es una peseta bien proporcionada —o cualquier otra fracción igualmente—). Arrojándola al aire un número suficientemente largo podemos acercarnos cuanto deseemos a $\frac{1}{2}$. Por ejemplo, en 10 tiradas quizá salgan 3 caras y 7 cruces. En 100 tiradas tendríamos 47 caras, en 1.000 podríamos tener 504, y así sucesivamente. Es decir, puede proponerse la hipótesis de que si las tiradas de la peseta se continúan indefinidamente, para cualquier número ϵ, por pequeño que sea,

habrá un número m de tiradas tal que para todo $n \geqslant m$ la proporción (o "frecuencia") de caras después de n tiradas es f_n tal que

$$|f_n - \tfrac{1}{2}| < \epsilon$$

Así, si se elige que ϵ sea 0,001 entonces

$$(\exists m)\,(n)\,[(n \geqslant m) \supset (\|f_n - \tfrac{1}{2}| < 0{,}0001)]$$

y en general tenemos

$$(\epsilon)\,(\ m)\,(n)\,[(n \geqslant m) \supset (|f_n - \tfrac{1}{2}\| < \epsilon]$$

Es importante el orden de los cuantificadores en la fórmula anterior. Obsérvese también que para cualquier elección respecto a lo que ϵ sea (tal como 0,0001) no *conocemos* ningún valor de m tal que para $n \geqslant m$ la diferencia entre f_n y $\frac{1}{2}$ sea menor que ϵ. Después de diez millones de tiradas es *posible,* aunque muy poco probable, que los próximos diez millones de tiradas sean todos "cara", en cuyo caso por supuesto $|f_n - \frac{1}{2}|$ no es inferior a 0,0001 para todo $n \geqslant 10.000.000$. La hipótesis afirma simplemente que *existe* tal valor de m.

Supóngase que para cualquier ϵ existe tal valor de m. Entonces podemos decir que la *frecuencia límite* de "caras" es $\frac{1}{2}$, es decir

$$\lim_{n \to \infty} f_n = \tfrac{1}{2}$$

La *teoría de la frecuencia* de la probabilidad define la probabilidad como frecuencia límite. Fórmulas del cálculo de probabilidades tales como $P(a \cdot b, c) = P(a, c) \times P(b, a \cdot c)$ siguen siendo verdaderas cuando se interpretan los diversos términos como frecuencias límite. Así la fórmula anterior afirma que la frecuencia límite de los sucesos que son a la vez a y b dentro de la clase de los sucesos c es igual a la frecuencia límite de los sucesos a dentro de la clase de los sucesos c multiplicada por la frecuencia límite de los sucesos b dentro de la clase de sucesos que son a la vez a y c.

Sería conveniente, a estas alturas, describir otra interpretación de la que es suceptible el cálculo de probabilidades. Hasta ahora hemos considerado las interpretaciones metalingüísticas (a menudo erróneamente denominadas "lógicas") y de la frecuencia. La próxima interpretación que se considerará es aquella en la que "$P(a, b)$", etc., se interpretan como *proporciones de posibilidades.* "$P(a, b)$" se interpreta como el número de casos posibles en que un suceso es a: Por ejemplo, sea b el número de caras de un dado (no hace falta que sea necesariamente uno que no haya sido "trampeado") y sea a el número de caras con numeración par. Entonces

$$P(a, b) = \frac{3}{6}$$

Supóngase que consideramos el lanzamiento simultáneo de dos dados. Entonces existen 36 posibles resultados distintos, a saber "uno" tirado con el primer dado y "uno" tirado con el segundo, "uno" con el primer dado y "dos" con el segundo, y así sucesivamente. Denominemos a estos 36 posibles resultados sucesos "*c*". Sea un suceso *a* una tirada de "uno" con el *primer* dado, y un suceso *b* una tirada de un número par con el *segundo* dado. Entonces es evidente que

$$P(a, c)=\frac{1}{6}$$

$$P(b, a \cdot c)=\frac{3}{6}$$

$$P(a \cdot b, c)=\frac{3}{36}$$

Es fácil comprobar que la fórmula

$$P(a \cdot b, c)=P(b, a \cdot c)\times P(a, c)$$

sigue siendo verdadera, así interpretada, como una fórmula acerca de posibilidades. La fórmula anterior, así interpretada, no nos dice *nada* acerca de la *probabilidad* de obtener ciertos resultados después de lanzar los dados, ya que la interpretación es válida exactamente igual ya sean los dados correctos o no. Esto a menudo no se aprecia debido al hecho de que las distintas posibilidades (por ejemplo cada uno de los 36 casos anteriores) son consideradas a menudo tácitamente "equiprobables" (es decir, de acuerdo con la interpretación de la frecuencia de la probabilidad), y entonces, desde luego, el cálculo de probabilidades sí que proporciona información acerca de las probabilidades (frecuencias) también.

Podemos considerar, entonces, un cálculo de probabilidades abstracto que es susceptible, cuando menos, de los tres tipos principales de interpretación siguientes: el metalingüístico, el de la frecuencia y la interpretación de la posibilidad. De ellos el de mayor interés para la filosofía de la ciencia es el de la interpretación de la frecuencia. Los enunciados dentro de la interpretación de la posibilidad por sí mismos no pueden decirnos nada acerca de lo que realmente ocurre, y la interpretación metalingüística es por sí misma divisible en varias subinterpretaciones. ¿Qué tipo de enunciado es el enunciado de que $P(a, b)=r$ siendo r algún número real entre 0 y 1? Quizás podría ser el enunciado de que sobre la evidencia de *b* resulta *racional creer* la proposición *a* con un grado de creencia igual a *r*, y podría existir algún modo de asignar el número *r* descubriendo cómo una persona racional, cuya única evidencia relevante fuese *b*, podría arriesgarse a mantener la verdad de *a*. En la mayoría de los casos, sin embargo, habrá dificultades en esto. ¿Podemos nosotros, de esta forma (o en cualquier otra) asignar probabilidad numérica *alguna* a la pro-

babilidad de que la teoría general de la relatividad sea verdadera? ¿Cómo podríamos ni siquiera imaginar lo que sería racional arriesgarse a mantener sobre su verdad? Más aún, el definir la probabilidad en términos de racionalidad está expuesto a la acusación de circularidad, especialmente si no se puede dar ninguna definición de "creencia racional" que no sea "creencia en lo que es más probable". Existe, sin embargo, una clase restringida de casos en la que la interpretación metalingüística opera perfectamente, porque opera a modo de parásito de la interpretación de la frecuencia. Es decir, sea K una clase de objetos, sea L una clase bien ordenada de objetos y sea r la frecuencia límite de los objetos que son L (con relación a la clase K). Entonces, si no se sabe nada acerca de algo sino que es miembro de K, podemos asignar inocuamente r como la probabilidad (metalingüística) de la proposición de que este objeto es un miembro de la clase L. Es decir, en algunos casos podemos asignar probabilidades metalingüísticas numéricas que son parásitos de probabilidades obtenidas sobre la base de frecuencias conocidas.

En este libro en ocasiones desearé hacer uso de enunciados sobre la probabilidad metalingüística que no parecen ser parásitos de enunciados acerca de frecuencias. Por ejemplo, desearé afirmar que, en igualdad de condiciones, la más sencilla de dos hipótesis tiene más probabilidades de ser verdadera. Debo admitir que no sé cómo justificar esta afirmación, por lo demás bastante plausible, o incluso manifestar con toda claridad lo que significa. Ciertamente no parece plausible apoyarla en la dudosa afirmación de que en términos generales las teorías sencillas han resultado ser verdaderas. (De ser así entonces esta afirmación metalingüística particular dependería de una afirmación metalingüística de la *frecuencia*.)

La interpretación de la frecuencia de la probabilidad y la interpretación de la tendencia

Se mantiene, normalmente, que las afirmaciones de probabilidad *dentro* de teorías científicas (en oposición a *acerca de* ellas) deben ser analizadas como afirmaciones acerca de frecuencias límite. K. R. Popper [8] ha manifestado recientemente el punto de vista de que deberían ser interpretadas más bien como tendencias de sistemas físicos. La diferencia entre estos dos puntos de vista puede establecerse del modo siguiente: los teóricos de la frecuencia desean analizar las probabilidades en términos de frecuencias observadas (como valores límites de éstas), mientras que, según la interpretación de la tendencia, los enunciados de probabilidad no pueden traducirse en enunciados acerca de frecuencias, aunque sean *comprobados por medio de* observaciones de frecuencias.

La principal objeción de Popper a la teoría de la frecuencia es la siguiente. Supóngase que tengamos un dado trampeado, de modo que después de una larga secuencia de tiradas la proporción de tiradas en las que sale "seis" es aproximadamente de $\frac{1}{4}$. Entonces habría que conjeturar

que la frecuencia límite es de aproximadamente $\frac{1}{4}$. Ahora considérese otra secuencia, que es la misma de antes (a saber, una repetición de duración indefinida de tiradas con el dado trampeado), pero que incluya, por ejemplo, cinco tiradas con un dado sin trampear. La frecuencia límite de esta nueva secuencia de tiradas sigue siendo $\frac{1}{4}$, ya que cinco tiradas con el otro dado no alteran la frecuencia a la larga. (Todavía aproximadamente $\frac{1}{4}$ resultarán "seis".) Sin embargo, con seguridad incluso, un teórico de la frecuencia mantendrá que para una tirada con un dado sin trampear la probabilidad de conseguir "seis" no es $\frac{1}{4}$, sino $\frac{1}{6}$. El teórico de la frecuencia dará como razón el que si el dado sin trampear fuese arrojado muchísimas veces la proporción de seises en su secuencia de tiradas se acercaría a $\frac{1}{6}$. Popper considera que al hacer tal modificación en su teoría el teórico de la frecuencia se ha pasado a la teoría de la tendencia. No estamos considerando secuencias reales (ya que evidentemente las pocas tiradas con el dado sin trampear son miembros perfectamente válidos de la secuencia de duración indefinida consistente en las pocas tiradas con el dado sin trampear junto la cantidad indefinida de tiradas con el dado trampeado) sino cálculos experimentales. *El tirar un número indefinido de veces un dado trampeado junto con cinco tiradas con el trampeado* constituye un cálculo experimental, mientras que *el tirar el dado no trampeado* constituye otro cálculo experimental. Por supuesto, que una tirada individual real puede ser considerada como perteneciente a más de un cálculo experimental.

Si aceptamos la descripción de las propiedades disposicionales que será ofrecida posteriormente en el capítulo 5, podemos identificar las tendencias con estados físicos reales. Así, la tendencia a que salga "seis" con un dado puede identificarse con (a) la distribución simétrica o asimétrica del material en el dado cúbico *junto con* la estructura neurofisiológica del jugador humano de dados, que explica por qué la velocidad angular que éste imparte al dado depende más bien del azar y no guarda relación de forma muy clara con la distancia de sus dedos de la mesa, y así sucesivamente. Aquí una tendencia es una subestructura de la disposición física que también habrá que describir. En la mecánica cuántica, sin embargo, no nos encontramos con tendencias posteriormente descriptibles (al menos si la mecánica cuántica de hoy en día está en lo cierto). Sin embargo, aunque no son posteriormente descriptibles deben ser consideradas como estados físicos objetivos. El propio Popper llegó a la teoría de la tendencia a través de consideraciones de la mecánica cuántica, y discutiremos algunas de ellas más adelante, en el capítulo 5.

Una gran ventaja de la teoría de la tendencia es que según ella una probabilidad es una faceta de un cálculo experimental que formulamos como hipótesis sobre la base (quizá) de frecuencias observadas, pero la teoría no *define* las probabilidades de este modo. La teoría de la frecuencia precisa de idealización o ficción. ¿Qué ocurriría en realidad si se diese un gran número de secuencias de tiradas con un dado? El dado llegaría a desgastarse y mucho antes de ello el jugador de dados se moriría. Por tanto, los

teóricos de la frecuencia tienen que hablar en realidad de lo que ocurriría en una secuencia infinita de tiradas que no tiene posibilidades de existir. Los teóricos de la tendencia pueden evitar este tipo de dificultad. Pueden analizar un solo caso como una tendencia de un cálculo experimental que aquél ejemplifica, y esta tendencia puede ser un estado físico real.

La relación entre la teoría de la tendencia y la teoría de la frecuencia es bastante semejante a la que se da entre 1) una teoría que identifica la vanidad con un estado neurofisiológico de la persona vanidosa (aunque quizá un estado desconocido) que *explica* por qué (por ejemplo) ésta lleva un chaleco amarillo y, 2), una explicación tal como la de Ryle, que intenta elucidar la vanidad simplemente en términos de enunciados hipotéticos acerca del comportamiento, tales como "si se cumplen tales y tales condiciones esta persona usa un chaleco de colores chillones". La dificultad de dar análisis plausible alguno de la vanidad siguiendo la línea de Ryle es bien conocida. Es difícil especificar las proposiciones hipotéticas que serían necesarias. Considérese el caso del hombre vanidoso que consigue ocultar su vanidad (y que quizá se envanece de no parecer vanidoso). Por otra parte, no hay dificultad alguna si la conducta vanidosa no se considera como *lo que constituye* la vanidad sino *lo que se explica* mediante la vanidad. De igual modo, el elemento de verdad en la teoría de la frecuencia de la probabilidad es que las hipótesis acerca de las probabilidades son comprobadas mediante frecuencias observadas, pero un cálculo puede presentar un resultado más o menos probable (tendencia de una magnitud mayor o menor a producir el resultado) se dé o no se dé de hecho tal resultado.

Hay un hecho importante que considerar acerca del uso de las frecuencias observadas a fin de comprobar los enunciados acerca de las probabilidades. Tanto los que abogan por la teoría de la tendencia como los teóricos de la frecuencia estarán de acuerdo en esto —en que estas pruebas no son concluyentes en el sentido en que lo puedan ser las pruebas de las generalizaciones universales. Es decir, supóngase que consideramos la generalización de que todos los cuervos son negros. Una sola observación de un cuervo azul, siempre que pudiésemos estar del todo seguros de que tal objeto debe ser clasificado como cuervo y de que es azul [8], serviría para refutar la hipótesis de que todos los cuervos son negros. Sin embargo, cinco tiradas sucesivas no refutarían de modo concluyente la hipótesis de que el dado no había sido trampeado (es decir, la probabilidad de que salga cada una de las caras es de un sexto). Rechazaríamos provisionalmente la hipótesis si estas cinco tiradas constituyesen toda la evidencia con la que contásemos acerca del dado, pero rechazaríamos nuestro rechazo si continuásemos tirando y encontrásemos (por ejemplo) un seis más, seis cincos, cinco cuatros, siete tres, tres doses y ocho unos

[8] Se presentan algunas dificultades con relación a esto que serán discutidas en el capítulo 6. No afectan a este apartado.

en las treinta tiradas siguientes. Desde luego, si manejamos largas secuencias de hechos podemos contar con base bastante sólida para rechazar hipótesis de probabilidad. La probabilidad de que salgan cinco seises con un dado no trampeado es sólo de $1/6^5$, y, por tanto, si estas tiradas constituyen todo nuestro conocimiento relevante, sería prudente que rechazásemos la hipótesis de que el dado no está trampeado. (Obsérvese aquí que tenemos un enunciado de probabilidad metalingüística que es parásito de un enunciado de tendencia o frecuencia.) Con todo, cinco seises sucesivos, o incluso cinco mil seises sucesivos, no son *lógicamente incompatibles* con la hipótesis de que el dado no está trampeado en el sentido en que "esto es un cuervo azul" es lógicamente incompatible con "todos los cuervos son negros".

Sugerencias respecto a lecturas posteriores

En este capítulo he seguido, en gran parte, el método y terminología de W. V. Quine, y su breve resumen de la lógica simbólica en la *Encyclopedia Americana* (1957 y ediciones posteriores), podría muy bien ser recomendado a los lectores con preferencia a la primera parte del presente capítulo. Ha sido convenientemente reimpreso en la obra de Quine, *Selected Logic Papers* (New York: Random House, 1966), págs. 37-51. Los *Methods of Logic* de Quine, segunda edición, revisada (New York: Holt, Rinehart and Winston, 1959) es un libro de texto particularmente iluminador y filosóficamente agudo sobre la moderna lógica sentencial y de predicados, con incursiones en la teoría de conjuntos. Este libro utiliza un método de "deducción natural" que presenta ventajas técnicas al facilitar la realización de pruebas, pero que es más difícil de captar que la presentación axiomática habitual. Para esto último véase D. Hilbert y W. Ackermann, *Principles of Mathematical Logic,* traducido por L. M. Hammond, G. G. Leckie y F. Steinhardt, editado con notas por R. E. Luce (New York: Chelsea, 1950). Un buen libro de texto sobre lógica es el de I. Copi, *Symbolic Logic,* segunda edición (New York: Macmillan, 1965). Un tratado a un nivel bastante avanzado y bastante detallado es el de Alonzo Church, *Introduction to Mathematical Logic* (Princeton: Princeton University Press, 1956). Sobre teoría de conjuntos véase W. V. Quine, *Set Theory and its Logic* (Cambridge: Harvard University Press, 1963). La cuestión de los fragmentos de la lógica de predicados para los que existen procedimientos de decisión es discutida por W. Ackermann, *Solvable Cases of the Decision Problem* (Amsterdam: North-Holland, 1954) y 2.ª y 3.ª edición, New York: Humanities Press, 1962). Sobre sintaxis, computabilidad y problemas de decisión véase S. C. Kleene, *Introduction to Metamathematics* (Princeton: D. Van Nostrand, 1952, traducción española: *Introducción a la Metamatemática,* ed. Tecnos, Madrid, 1974); Martin Davis, *Computability and Unsolvability* (New York: McGraw-Hill, 1958); y R. Smullyan, *Theory of Formal Systems,* edición revisada (Princeton: Princeton University Press, 1961). Todos estos libros precisan poco o ningún conocimiento matemático previo, pero sí requieren capacidad para la matemática por parte del lector. El libro de Smullyan es particularmente conciso y hace uso de métodos muy elegantes y simplificadores, pero el método de Martin Davis, por medio de la teoría de las máquinas Turing, quizá resulte más fácil intuitivamente para aquellos lectores que tienen más facilidad para pensar acerca de estas cuestiones en términos de lo que hacen las máquinas computadoras que en términos de sistemas formales abstractos. (Aunque desde un punto de vista

matemático estos dos métodos vienen a resultar muy semejantes.) Métodos más asequibles de algunas de estas cuestiones pueden encontrarse en A. M. Turing, "Solvable and Unsolvable Problems", *Penguin Science News,* número 31 (1954), 7-23; E. Nagel y J. R. Newman, *Gödel's Proof* (New York: New York University Press, 1958, traducción española: *El teorema de Gödel,* ed. Tecnos, Madrid, 1970) [en unión del cual consúltese la revisión crítica de Hilary Putnam en *Philosophy of Science,* XXVII (1960),205-207], y W. V. Quine, "Foundations of Mathematics", *Scientific American* (septiembre 1964), 112-127, reimpreso en su obra *The Ways of Paradox and Other Essays* (New York: Random House, 1966), págs. 24-34. Otro valioso artículo es el de Barkley Rosser, "An informal exposition of proofs of Gödel's theorems and Church's theorem", *Journal of Symbolic Logic,* IV (1939), 53-60. Argumentos en contra de las lógicas intensionales pueden verse en W. V. Quine, *Word and Object* (Cambridge: Massachusetts Institute of Technology Press, 1960), capítulo 6, especialmente la sección 41; y "Three Grades of Modal Involvement", *Proceedings of XIth International Congress of Philosophy,* XIV (Bruselas, 1953), 65-81, reimpreso en la obra de Quine, *The Ways of Paradox,* págs. 165-174. Véase también el trabajo de Quine, "Reply to Professor Marcus", reimpreso en *The Ways of Paradox,* págs. 175-182. Esto, junto con el trabajo de la profesora Marcus y una interesante discusión entre Quine, Marcus y otros, apareció originariamente en *Synthese,* XX (1961) y en M. W. Wartofsky (ed.), *Boston Studies in the Philosophy of Science,* vol. I (Dordrecht, Holland: D. Reidel, 1963, y New York: Humanities Press, 1963).

Sobre la probabilidad véase E. Nagel, *Principles of the Theory of Probability, International Encyclopedia of Unified Science,* vol. I, número 6 (Chicago: University of Chicago Press, 1938); William Kneale, *Probability and Induction* (Oxford: Clarendon Press, 1949, y New York: Oxford University Press, 1949); y G. H. von Wright, *A Treatise on Induction and Probability* (London: Routledge & Kegan Paul, 1951). Una relación de la teoría de la frecuencia de la probabilidad, escrita por uno de sus principales defensores, y en un lenguaje accesible al lector no matemático, es la de R. von Mises, *Probability, Statistics and Truth,* segunda edición (London: Allen & Unwin, 1957). Una teoría de la frecuencia de la probabilidad aparece desarrollada en la obra de K. R. Popper, *The Logic of Scientific Discovery* (London: Hutchinson, 1959, y en rústica, edición revisada, New York: Harper and Row, 1965, traducción española: *La lógica de la investigación científica,* ed. Tecnos, Madrid 1962), que es una traducción con apéndices adicionales y notas al pie de la obra de Popper, *Logik der Forschung* (Viena: Julius Springer, 1934). Desde entonces, como se observaba en el texto, Popper ha vuelto a su "interpretación de la tendencia" de la probabilidad, como por ejemplo en su artículo "The Propensity Interpretation of Probability", *British Journal for the Philosophy of Science,* X (1959), 25-42. C. G. Hempel en sus *Aspects of Scientific Explanation* formula una explicación disposicional de la probabilidad que es semejante a la teoría de la tendencia de Popper y mantiene que Rudolf Carnap, en su "Inductive Logic and Science", *Proceedings of the American Academy of Arts and Sciences,* LXXX (1951-54), 187-197, sostiene de modo semejante que la probabilidad estadística de conseguir un as con un dado determinado es un estado físico ("el estado de probabilidad") del dado, del cual son síntomas las frecuencias relativas. Carnap ha escrito extensamente sobre la probabilidad, principalmente con vistas a sus repercusiones en el problema filosófico de la inducción y, por tanto, el tono del trabajo es a menudo bastante especializado. Véase, por ejemplo, su *Logical Foundations of Probability,* segunda edición (Chicago: University of Chicago Press, 1962) y el resumen de su postura en las págs. 966-979 de su "Replies and Expositions", en P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (La Salle, III: Open Court, 1963). A un nivel más asequible véase el artículo de Carnap, "What is Probability?", *Scientific American* (septiembre 1953), 128-138. Otro útil trabajo es el de I. J. Good, "Kinds of Probability", *Science,* CXXIX (1959), 443-447.

3. EXPLICACION, LEYES Y TEORIAS

Observaciones preliminares acerca de la explicación

Estamos ahora preparados ya para empezar a hablar acerca de la ciencia. Es innegable que una de las principales cosas que esperamos obtener de la ciencia es la capacidad de *explicar* diversos tipos de cuestiones. Consideremos el caso del muchacho emocionado con la física que aprende en el bachillerato, ya que le permite comprender bastante bien, cómo funciona su aparato de radio. Quizá posee una radio con transistor. Pudiera ser que conozca una explicación particular que le permita saber únicamente que los transistores realizan las mismas funciones que las válvulas de un modelo más antiguo. Ahora su asombro ya no es acerca del funcionamiento de su radio en conjunto sino de los transistores que están dentro de ella, y esto le conduce a la moderna física del estado sólido. Revistas divulgadoras o semidivulgadoras como la *Scientific American,* se ocupan, meritoriamente, de este tipo de curiosidad. ¿Cuál debe ser la sensación de un ama de casa (o tal vez de un crítico literario) que no tiene idea *alguna* acerca de cómo funciona la radio, cómo el frigorífico que está en la cocina enfría lo que tiene dentro, o por qué la expulsión de gases calientes por los extremos traseros del motor de reacción hace que un avión se acelere hacia adelante? O, en la esfera "natural" (en oposición a la tecnológica), considérense cuestiones tales como por qué aparece el arco iris, por qué la cola de los cometas se separa del Sol, cómo se reproducen las plantas, por qué el polen hace estornudar a algunas personas y demás. Todas éstas son preguntas cuya *formulación* precisa poca preparación científica, aunque quizá su *contestación* requiera bastante preparación. Los científicos se ocupan, a menudo, de contestar preguntas fácilmente formuladas, aunque, por supuesto, también se ocupan con frecuencia de preguntas que ni siquiera pueden ser formuladas en términos vulgares. Considérese la de "¿Cómo disuelve el lisocimo los muco-polisacáridos?" [1]. Se necesita mucha preparación científica antes de poder comprender lo que significa tal pregunta.

Dejando completamente aparte los beneficios prácticos que nos proporciona la ciencia (y estos beneficios prácticos se derivan todos en última instancia del hecho de que la ciencia nos permite *predecir* lo que ocurrirá en diversas combinaciones de circunstancias) parece innegable que gran parte del atractivo intelectual de la ciencia se deriva de su capacidad para ayudarnos a contestar a muchos de los "porqués" y "cómos" de los que dimos muestras en el parágrafo anterior. Se dice a veces, por supuesto, que la ciencia nos permite responder al "cómo", pero no al

"por qué". Considerada literalmente tal afirmación resulta evidentemente falsa, ya que muchas de las cuestiones a las que la ciencia responde se formulan con toda naturalidad con un "¿por qué...?". Considérese la pregunta que se mencionó en el parágrafo anterior de "¿por qué la cola de los cometas se separa del Sol?". (Respuesta: debido a la presión de la radiación solar sobre las pequeñas partículas que constituyen la cola.) Esta cuestión se formula mucho más naturalmente con un "por qué" que con un "cómo", como por ejemplo: "¿Cómo se separan del Sol las colas de los cometas?"

"¿Cómo se separan del Sol las colas de los cometas?", no es una cuestión muy clara, la cual es más natural que se interprete no como la petición de una explicación sino como la solicitud de una descripción más detallada del modo en que la cola de los cometas se separa del Sol (por ejemplo, si mide muchos millones de metros). A veces "por qué" y "cómo" son simplemente intercambiables. No parece ser que haya diferencia alguna entre preguntar: "¿Por qué la expulsión de gases calientes por los extremos traseros del motor de reacción hace que un avión se acelere hacia adelante?" y la pregunta: "¿Cómo la expulsión de gases calientes por los extremos traseros del motor de reacción hace que un avión se acelere hacia adelante?" Ambas formas de la pregunta se contestarían igualmente, y de modo natural, mediante una referencia a la tercera ley del movimiento de Newton o al principio de la conservación de la fuerza. También "¿por qué el lisocimo disuelve los muco-polisacáridos?" normalmente se considera que significa lo mismo que "¿cómo...?", aunque supongo que podría considerarse de otro modo; por ejemplo, como una pregunta acerca del papel que desempeña la disolución de los polisacáridos en el funcionamiento de un organismo. En algunos casos es más natural la interpretación teleológica o funcional del "por qué". "¿Por qué tienen dos estómagos las vacas?", podría contestarse diciendo: "a fin de digerir la hierba", o algo por el estilo. Otra explicación de este tipo tendría que hacerse a través de la teoría de la evolución mediante la selección natural: el tener dos estómagos (y por tanto ser capaz de digerir la hierba) es útil (en ciertas circunstancias) para la supervivencia. Tendremos que discutir las explicaciones teleológicas en el capítulo 4, pero en el contexto de la teoría de la evolución son perfectamente compatibles con las explicaciones científicas ordinarias. La idea de que la ciencia no puede dilucidar los "porqué" con toda seguridad procede de la interpretación de los "porqué" como cuestiones acerca de las intenciones de Dios, que creó el mundo. Hemos visto, sin embargo, que las preguntas con "por qué" son también, con frecuencia, preguntas con "cómo". Incluso cuando son teleológicas (y por tanto difieren de las preguntas con "cómo" correspondientes) no hay necesidad de nada extracientífico o teológico para contestarlas, ya que la aparente referencia a la intención puede desaparecer cuando se juzga la explicación a la luz de la teoría de la evolución mediante la selección natural.

¿Qué es la explicación?

En la sección precedente hemos considerado algunos ejemplos de solicitud de explicación. ¿Podemos dar una caracterización general de la explicación? Si con esto se quiere decir, "¿podemos dar una definición general de la palabra "explicación"?", nos tropezamos con la dificultad de que la palabra se usa en variedad de contextos notablemente diferentes, algunos de los cuales nada tienen que ver, en absoluto, con el concepto de explicación científica.

Así, un matemático puede comenzar su libro con una explicación de sus símbolos, un escolar puede explicar el significado de un texto y una azafata puede explicar el modo de utilizar los cinturones de seguridad y las mascarillas de oxígeno. Probablemente lo que hay de común entre estos tres usos de la palabra "explicación" y el uso científico es que en todos los casos anteriores al parecer se elimina algún tipo de enigma. En algunos casos, por supuesto, quizá no haya ningún enigma que descubrir: la explicación de un texto puede ser dado por un examinado que no presupone ignorancia alguna por parte de sus examinadores (o al menos ninguna ignorancia acerca del texto), y la azafata puede estar llevando a cabo lo regulado, aunque sepa que los viajeros son ya veteranos y conocen las instrucciones en cuestión.

En los mencionados contextos no científicos el verbo "explicar" es un predicado triádico: es decir, *alguien* explica *algo* a *alguien*. La palabra "explicar" también se usa a menudo, en este sentido, en contextos científicos, como en "Newton explicó a los lectores de los *Principia* por qué los planetas se mueven en órbitas aproximadamente elípticas". En este sentido "explica" opera más bien como "demuestra". Podemos decir que un matemático "demuestra un teorema" *simpliciter* (es decir, lo deduce de acuerdo con reglas correctas de la lógica a partir de sus axiomas), pero también podemos decir que se lo demuestra a una determinada persona, en cuyo caso se sugiere que él no sólo deriva correctamente el teorema sino que también hace que otra persona comprenda y acepte tal derivación. En metamatemática, por supuesto, nos abstraemos por encima de este aspecto pragmático de la prueba (es decir, por encima de las cuestiones de los *efectos* de las pruebas en el auditorio) y consideramos las pruebas únicamente desde el punto de vista de si son o no deducciones correctas a partir de los axiomas. No consideramos ni al que proporciona la prueba ni al que la recibe, sino simplemente las sentencias de la prueba como estructuras abstractas. Ahora bien, al discutir la explicación algunos autores (por ejemplo C. G. Hempel [2]) prefieren discutir el concepto de explicación abstrayéndose de los aspectos pragmáticos cuanto sea posible: es decir, se concentran en un concepto de la explicación que se define en términos de su forma lógica (y en un concepto de la explicación correcta que depende también de la *verdad* de sus premisas). Es decir, intentan dar una definición de la explicación (que si la explicación se

expresase en un lenguaje formal) pertenecería a la sintaxis y a la semántica. Por otra parte, algunos escritores, como Michael Scriven [3], acentúan los aspectos contextuales y pragmáticos de la explicación (tales como la solución de enigmas) y mantienen que una definición que sea una abstracción por encima de tales aspectos no parece que sea fructífera. Es decir, estos últimos filósofos mantendrían que cuando nos abstraemos por encima de los aspectos contextuales y pragmáticos lo que queda no tiene gran importancia, y que una teoría que se abstrae por encima de esos aspectos no puede compararse con una disciplina de tal interés como la teoría de la prueba en metamatemáticas. De pasada debe observarse que ni siquiera la teoría de la prueba en metamatemáticas puede abstraerse del todo por encima de las consideraciones pragmáticas, por cuanto se dan disputas entre distintos filósofos de las matemáticas con relación a qué es lo que constituye una prueba *aceptable*; algunos de ellos (matemáticos "intuicionistas" o "constructivistas") no aceptarán formas de pruebas que son aceptables para otros (los matemáticos "clásicos"). De todos modos, la teoría de la prueba presenta tal riqueza de logros técnicos (tales como las pruebas de la completitud y la incompletitud, decidibilidad e indecidibilidad) que las discusiones acerca de los aspectos pragmáticos quedan muy oscurecidas. Debe observarse que aunque, en la práctica, los defensores de los métodos formalista y pragmático de discusión de la explicación disputan a veces entre sí no se da, necesariamente, incompatibilidad alguna entre los dos métodos. En gran parte puede ser sólo cuestión de lo que distintos filósofos consideren que es lo más *interesante* de la explicación.

Consideremos algo de lo que puede decirse acerca de la explicación desde el punto de vista más formal. Hempel [4] ha distinguido dos tipos principales de explicación científica, que denomina, respectivamente, *deductivo-nomológica* e *inductivo-estadística.*

El modelo de explicación deductivo-nomológica

Supóngase que se hace girar una peonza con una velocidad angular ω sobre su eje, de modo que se mueva suavemente en vertical dando vueltas sobre el punto donde su punta se apoya en el suelo. ¿Por qué no se cae de lado como ocurriría si no se le hiciese girar muy de prisa? Aquí nos encontramos con un hecho particular que hay que explicar, a saber, el movimiento de esta peonza particular. Representemos la sentencia que expresa este hecho particular mediante "E". Entonces hay sentencias que describen la forma exacta de la peonza, su composición (supongamos) homogénea (por lo menos en un grado bastante aproximado), su masa y su velocidad angular de rotación sobre su eje y su ángulo con la verticular cuando se hizo girar. Sean estas sentencias C_1, C_2, C_3, ..., C_m. Estas expresan lo que se denomina comúnmente "condiciones iniciales", aunque la palabra "inicial" no debe considerarse que implica que las condiciones

preceden temporalmente al fenómeno que se ha de explicar. En nuestro caso sí lo preceden, pero no tiene que ser así necesariamente en otros casos. Supóngase que lo que tuviéramos que explicar, basándonos en los datos acerca del movimiento actual de la peonza, fuese por qué no se caía hacia un lado con *un poco* de anterioridad, las sentencias que expresasen su estado de movimiento *actual* seguirían siendo consideradas como condiciones "iniciales". "Inicial" aquí significa "inicial respecto al problema que tenemos entre manos" más que "inicial en el tiempo".

Además de C_1, C_2, C_3, ..., C_m necesitamos leyes L_1, L_2, ..., L_n, que en nuestro ejemplo de la peonza serán las leyes de la mecánica newtoniana. Entonces es posible construir un argumento

$$\begin{array}{l} C_1, C_2, \ldots, C_m \\ L_1, L_2, \ldots, L_n \\ \hline \text{Por consiguiente } E \end{array}$$

Este es un ejemplo del modelo de explicación deductiva nomológica. Al parecer sea o no sea necesaria o suficiente para la explicación la posesión de esta forma, es un modelo de explicación muy característico. Para hacer formal la definición tenemos que caracterizar a C_1, C_2, ..., C_m y a L_1, L_2, ..., L_n formalmente. Ahora bien, C_1, C_2, ..., C_m son posiciones claramente particulares: se refieren a acontecimientos particulares o estados de hechos en regiones particulares del espacio y del tiempo. Podemos considerarlos como proposiciones que contienen nombres o como descripciones definidas de entidades particulares. Los enunciados legales, L_1, L_2, ..., L_n, por otra parte, son proposiciones que comienzan por cuantificadores universales y que no contienen nombres ni descripciones definidas. Es cierto que algunas proposiciones que hacen referencia a entidades particulares o regiones del espacio-tiempo se denominan a veces leyes; considérese, por ejemplo, la "ley" de que los cuerpos situados cerca de la superficie de la Tierra caen con una aceleración de aproximadamente 981 cm/sec^2. Sin embargo, podemos sustituir la referencia a la Tierra por "sobre la superficie de una esfera de tal y tal masa y radio", e insertar el enunciado adicional de que la Tierra es un cuerpo con tal masa y radio. De este modo separamos la proposición original en dos partes, una parte L y una parte C. A buen seguro que Galileo no podría haber hecho esto, en parte porque carecía de los datos geográficos necesarios, pero yo creo que hoy en día cualquier "ley" que contenga una referencia a una región particular del espacio-tiempo, y que sea de mucho interés para la filosofía de la ciencia, puede dividirse en una auténtica parte L y una auténtica parte C.

Existe, sin embargo, una objeción más grave a nuestra caracterización de las leyes como proposiciones completamente generales. Posiblemente no exista ningún otro cuerpo en el universo distinto que la Tierra con la misma masa, el mismo radio y la misma distancia de la estrella de la que

es planeta. En cuyo caso *cualquier* enunciado que mencione a la Tierra será verdadero si y sólo si un enunciado correspondiente perfectamente universal que en lugar del nombre "la tierra" contenga la cláusula "todos los cuerpos que tienen tal y tal masa, radio, distancia de su estrella" es también verdadero. Y si los predicados "que tiene tal y tal masa, etc.", no son suficientes, podemos añadir otros más, tales como "que es un planeta de una estrella con tal y tal espectro". En resumidas cuentas, si existe uno y sólo un objeto en el universo respecto al cual los predicados P_1, P_2, ..., P_k, sean verdaderos siempre podemos sustituir la proposición particular:

$$F\,[(_1x)\,(P_1x \cdot P_2x \ldots P_kx)]$$

por la *general*

$$(x)\,[(P_1x \cdot P_2x \ldots P_kx) \supset Fx]$$

siempre que nuestro discurso esté en un lenguaje puramente extensional. Existen dos caminos a seguir por el filósofo. Podría considerar la proposición mencionada:

$$(x)\,[(P_1x \cdot P_2x \ldots P_kx) \supset Fx]$$

y sus semejantes como leyes perfectamente válidas, y podría, por consiguiente, sustituir las condiciones iniciales C_1, C_2, ..., C_m, de esta forma, y de modo semejante podría sustituir la descripción particular E del acontecimiento que se ha de explicar por una proposición correspondiente perfectamente general. Tendríamos entonces el caso de la explicación de una ley mediante otras leyes. También podría renunciar al intento de definir leyes de modo puramente formal y consideraría una proposición general tal como la nuestra:

$$(x)\,[(P_1x \cdot P_2x \ldots P_kx) \supset Fx]$$

citada anteriormente, como una generalización más *accidental*. Intentaré aclarar esto mediante un ejemplo.

Consideremos una estrella ficticia, Delta Chaeropi (en la constelación ficticia Bandicoot). Esta estrella tiene seis planetas, y cada planeta tiene un número de satélites que da la coincidencia que es un número primo. Entonces la proposición:

1) (x) (x es un planeta de Delta Chaeropi $\supset$ el número de satélites de x es primo)

es verdadera respecto a todos los objetos del universo. Porque o bien un objeto no es un planeta de Delta Chaeropi o de lo contrario tiene un número primo de satélites. Con todo, ¿deseamos considerar 1) como ley de la naturaleza? Con toda seguridad es mera coincidencia el que los pla-

netas de Delta Chaeropi tengan un número primo de satélites. (Por supuesto que quizá no esté del todo claro, a estas alturas, lo que significa, si es que significa algo, el decir "mera coincidencia", pero aceptémoslo provisionalmente con la esperanza de aclararlo dentro de poco.)

Un científico tal vez negase que nuestra proposición 1) sea una ley de la naturaleza, simplemente basándose en que contiene el nombre propio "Delta Chaeropi". Sin embargo, puede existir un conjunto de predicados perfectamente generales (es decir, que no contengan ellos mismos nombres propios o palabras indicadoras como "aquí" y "ahora" [1]) que sirvan conjuntamente sólo para Delta Chaeropi. Supóngase que Delta Chaeropi es la única estrella del universo respecto a la cual sea verdadero un determinado conjunto finito de predicados "P_1", "P_2", "P_3", ..., "P_k". (Obsérvese que si el universo es infinito sería mucho suponer el que exista tal conjunto de predicados, pero, sin embargo, supongamos que tal supuesto es correcto.) Si 1) es verdadera entonces también lo es la proposición

2) $(x)\,[(P_1x \cdot P_2x \cdot P_3x \ldots P_kx) \supset$
x tiene un número primo de satélites]

No podemos negarle a 2) el "status" de ley de la naturaleza basándonos en que contiene algún nombre o alguna expresión de referencia única, ya que 2) no contiene ninguna expresión tal. Formalmente es tan general como cualquier ley de la naturaleza. Y, sin embargo, intuitivamente quizá sintamos la seguridad de que no es una ley sino simplemente una generalización accidentalmente verdadera.

Existe otra razón por la que puede fallar el intento de distinguir las leyes de la generalización como 1): nuestra distinción entre predicados perfectamente generales y predicados que contengan nombres propios o referencias a puntos particulares del espacio y el tiempo no es, en modo alguno, clara. Nelson Goodman ha definido el predicado "grue" con el significado de "verde con anterioridad al año 2.000 A. D. y azul con posterioridad", y a "bleen" lo la definido como "azul ("blue") con anterioridad al año 2.000 A. D. y verde ("green") con posterioridad" [5]. Si se dice que "grue" y "bleen" no son predicados perfectamente generales debido a que su definición hace referencia al año 2.000 A. D., puede replicarse (como lo hace Goodman) que "blue" y "green" no son perfectamente generales ya que pueden definirse como "'bleen' con anterioridad al año 2.000 A. D. y 'grue' con posterioridad" y como "'grue' con anterioridad al año 2.000 y 'bleen' con posterioridad". Intuitivamente vemos que "blue" y "green" son predicados más generales que "grue" y "bleen", y quizá podamos justificar esta intuición indicando cuánto más sencillo es enseñar a un organismo humano a aplicar los predicados "blue" y

[1] Por consiguiente "... a menos de 100 años luz del Sol" y "... a menos de 100 años luz de aquí" no se considerarán como predicados perfectamente generales.

* En inglés "green". *(Nota de la Traductora.)*

** En inglés "blue". *(Nota de la Traductora.)*

"green" que enseñarle a aplicar los predicados "grue" y "bleen". Lo que sí muestra el ejemplo, sin embargo, es que no puede establecerse la distinción de un modo puramente formal. De ser así, se seguiría evidentemente que no podemos distinguir leyes de la naturaleza de auténticas generalizaciones que impliquen una referencia a una región particular del espacio-tiempo de un modo puramente formal.

Una de las formas en que los filósofos han intentado distinguir las leyes de la naturaleza de las generalizaciones accidentalmente verdaderas ha sido diciendo que las leyes de la naturaleza amparan a los condicionales contrafácticos, mientras que las generalizaciones accidentales no los amparan. De este modo, supóngase que en la tecnología espacial del futuro los astrónomos son capaces de determinar que Delta Chaeropi tiene planetas y que estos planetas tienen satélites y que el número de los satélites de cada planeta resulta ser primo. (No hace falta decir que tales descubrimientos están por completo fuera del alcance de los astrónomos de hoy en día.) Los astrónomos no dirían que *si* Delta Chaeropi *tuviese* todavía otro planeta más, entonces el número de satélites de este planeta sería primo. Es decir, podríamos decir que nuestra proposición 1) no es una ley ya que no ampara a un condicional contrafáctico. Ahora bien, el condicional contrafáctico es, por supuesto, distinto del condicional ordinario "$\supset$". Lo más plausible es que realicemos un análisis metalingüístico del condicional contrafáctico. Decir "Si Delta Chaeropi hubiera tenido todavía otro planeta más, entonces los satélites de este nuevo planeta habrían sido primos en número" es decir que la sentencia "Los satélites del planeta $(n+1)$ de Delta Chaeropi son primos en número" es deducible de "Delta Chaeropi tiene $n+1$ planetas" (aquí estoy considerando que de hecho Delta Chaeropi tenga no $n+1$ sino n planetas), junto con determinados supuestos implícitos como base, por supuesto, tales como algunas leyes físicas de aceptación general. Es evidente, sin embargo, que no podría hacerse esto a menos que la proposición de que los satélites del planeta $(n+1)$ son primos en número fuese deducida de la ley de que si cualquier objeto es una estrella que se conforme a la descripción de Delta Chaeropi, entonces debe tener planetas cada uno de los cuales tiene un número primo de satélites. Es decir, que si nosotros intentamos determinar la diferencia entre las generalizaciones legales y las accidentalmente universales no podemos salir de un círculo vicioso, ya que ahora resultaría que los propios condicionales contrafácticos tienen que ser entendidos con referencia a la distinción entre generalizaciones legales y accidentales.

Esta circularidad ha sido observada por A. J. Ayer, quien agrega que su opinión personal es que la diferencia entre los dos tipos de generalizaciones (la generalización legal y la accidental) radica "no tanto en la diferencia de su contenido sino en nuestra actitud hacia ellas" [6]. Esta diferencia de actitud depende, en parte, de la diferencia de los papeles que desempeñan los dos tipos de generalización en las teorías explicatorias. En efecto, parece ser que consideramos que algo es ley si es o bien parte

de una teoría o si es algo que calculamos que algún día se integrará en una teoría. El científico calcula esto basándose en lo que siente en general hacia la materia de que se ocupa y no mediante criterios formales. Así, la mayoría de los científicos no considerarían significativo, ni más que una simple coincidencia, el que todos los planetas de una estrella tuviesen satélites cuyos números fuesen primos, al igual que, me imagino, la mayoría de los astrónomos no considerarán la denominada ley de Bode relativa a las distancias aproximadas de los planetas respecto al Sol como ley, incluso descontando la consideración de que hace referencia explícitamente a una entidad particular, el Sol[2]. Es decir, los astrónomos probablemente verían de un modo intuitivo, basándose en sus conocimientos de teoría física, que la ley de Bode no es más que una interesante coincidencia y que es improbable que llegue a integrarse en ningún sistema teórico[3]. Más aún, opino que los astrónomos seguirían pensando lo mismo aunque no se diese el caso de que algunos planetas diverjan de modo bastante notorio de los requisitos de la "ley" de Bode.

Explicación estadística

Hempel distingue dos tipos de explicación estadística. Al primero lo denomina explicación "Deductivo-Estadística". Esta es muy importante, debido, especialmente, a que de acuerdo con la mecánica cuántica las leyes fundamentales de la naturaleza son estadísticas. Según este modelo las leyes estadísticas se deducen de un conjunto de premisas que incluyen por lo menos una ley estadística. Si seguimos a Popper al interpretar los enunciados de probabilidad como enunciados acerca de tendencias objetivas más que como enunciados acerca de frecuencias límite, entonces el modelo deductivo-estadístico quedará reducido al deductivo-nomológico, con al menos una ley en las premisas que sea el enunciado de una ley acerca de tendencias.

La forma deductiva-nomológica se diferencia de la forma que Hempel denomina "Inductivo-Estadística"[4] en un aspecto más importante. Consi-

[2] Ley de Bode: Tómense los números 0, 3, 6, 12, 24, 48, 96, añádase 4 a cada uno de estos términos y divídase entre 10, obteniéndose así, 0,4, 0,7, 1,0, 1,6, 2,8, 5,2, 10,0... Estas cifras corresponden aproximadamente a las distancias medias a partir del Sol de los planetas Mercurio, Venus, la Tierra, Marte, Júpiter, Saturno (las verdaderas distancias en unidades astronómicas son 0,39, 0,72, 1,00, 1,5, 5,2, 9,54, si permitimos la separación en la que 2,8 viene a corresponder a los planetas menores, que ocupan una región entre Marte y Júpiter). Urano (descubierto después de la formulación de la "ley") también encaja perfectamente bien (distancia predicha 19,6, distancia real 19,19), pero Neptuno y Plutón divergen notoriamente (distancias predichas 38,8 y 77,2, respectivamente, distancias reales 30,07 y 39,46).

[3] He hablado de "integrarse en" más que de "deducirse de" ya que en toda teoría debe haber algunas leyes últimas no deducibles de otras. No obstante, podemos seguir diciendo que las leyes últimas se integran en la teoría de la que forman una parte esencial.

[4] Yo personalmente considero que esta teoría puede llevar a error, ya que la explicación de Hempel puede ser aceptada desde luego por aquellos que desean

dérese el siguiente ejemplo: Una proporción muy elevada de seres humanos pesan menos de 250 libras. Por tanto, podemos argumentar del siguiente modo:

La mayoría de los seres humanos pesan menos de 250 libras.
X es un ser humano.
..... [es muy probable que se dé el caso de que]
X pesa menos de 250 libras.

Hempel ha expresado el modelo en la forma acabada de indicar, en la que las palabras "es muy probable que se dé el caso de que" (o algo muy semejante) se insertan entre las premisas y la conclusión. Es evidente que al leer lo anterior debemos pronunciar el "es muy probable que se dé el caso de que" en un susurro, ya que de oírse serían necesarias las comillas abarcando las premisas y la conclusión, y el "argumento" anterior habría de considerarse no como un argumento sino como un *enunciado* metalingüístico. (Véanse las indicaciones de las páginas 36-7 del capítulo anterior.) Pero lo que Hempel quiere decir es que en el argumento anterior la relación entre las premisas y la conclusión no es la de la implicación sino la de "apoyo inductivo" o "confirmación". Esto equivale a la siguiente diferencia importante: si un argumento deductivo es bueno, continuará siendo bueno al añadírsele nuevas premisas adicionales. Esto no es así en el modelo inductivo-estadístico. Supóngase que contamos con la información adicional de que *X*, cuyo peso ha de explicarse, es una de las esposas de Rumanika, el rey de Karagwe en Africa Oriental, hace cien años. Rumanika consideraba que la mucha gordura era un ingrediente de la belleza femenina. Las reinas tenían que sorber leche mediante una pajilla prácticamente siempre que estaban despiertas, y se les azotaba si dejaban de chupar durante mucho tiempo [7]. Como resultado de ello les llegó a resultar bastante difícil el moverse. Pues bien, si sustituimos "*X* es un ser humano" por "*X* es un ser humano que es reina de Karagwe hace cien años" en el argumento anterior, nuestras premisas no hacen que sea probable que *X* pese menos de 250 libras, ya que la premisa adicional apoya fuertemente la hipótesis de que *X* pese por lo menos las 250 libras.

Supóngase, pues, que como explicación del hecho de que *X* pese menos de 250 libras se nos dice que la mayoría de los seres humanos pesan menos de 250 libras y que *X* es un ser humano, y por tanto es de esperar que *X* pese menos de 250 libras. Hay que admitir que esto no sirve mucho como explicación, pero según Hempel sería un determinado tipo de explicación. Ahora bien, observamos que para determinar su valor tenemos que fijarnos en algo exterior al modelo de argumentación de la propia explicación, de modo distinto a como hacemos en el caso del modelo de explicación deductiva. Tenemos que asegurarnos de que se satisface un

liberar a la ciencia de una explicación basada en el concepto del argumento inductivo. Véase el capítulo 6.

requisito *epistemológico,* a saber, que hemos hecho uso de toda la información relevante disponible. No aceptaríamos la explicación, quizás, si supiésemos que *X* era una reina de Karagwe (aunque podríamos aceptarla si supiésemos que *X* era una reina de Karagwe con alergia a la leche).

La moderna mecánica cuántica proporciona un campo en el que a veces podemos tener la completa seguridad de que se cumple el requisito epistemológico. Las leyes últimas de la mecánica cuántica son probabilísticas, y por tanto a menos que la mecánica cuántica (tal como ahora existe) sea falsa, se dan situaciones en las que es posible que haya evidencia adicional que pueda ser relevante. Considérese un electrón que pase a través de un cristal y se adhiera a una pantalla en una región *R*. El conocimiento completo de las condiciones antecedentes (el estado del electrón con anterioridad a su paso a través del cristal) nos permite determinar únicamente que existe tal y tal probabilidad de que el electrón se adhiera a la pantalla en la región *R*. No es posible, en este caso, ninguna evidencia adicional análoga a la de "*X* es una reina de Karagwe".

Ampliando el uso de una frase de William Dray [8], Hempel utiliza el término "modelo legal de cobertura" para abarcar tanto el modelo deductivo-nomológico como el inductivo-estadístico de explicación de hechos particulares.

Explicación incompleta

Hempel ha introducido dos nuevos conceptos que son particularmente importantes en relación con lo que estamos tratando. Son los de *explicación parcial* y *esbozo de explicación* [9]. Supóngase que se nos pide que expliquemos el hecho de que parte de un satélite de la Tierra cayese en Australia Central. Podríamos decir que el satélite no estaba completamente fuera de las regiones enrarecidas de la atmósfera terrestre, y que la fricción del paso del satélite a través de la atmósfera produjo como resultado una pérdida de energía cinética, y por tanto el satélite cayó a tierra. Esto podría considerarse como una explicación moderadamente satisfactoria, pero obsérvese que es una explicación de la caída en tierra del satélite (al menos si no está completamente vaporizado durante su caída a través de la atmósfera): no explica por qué el satélite cayó en Australia Central. Para explicar esto se necesitarían muchos más conocimientos acerca de la órbita exacta del satélite, cuándo se puso en órbita, la posición de Australia Central en la superficie terrestre y probablemente otros diversos hechos relevantes. Podemos decir que la explicación original era (o podría fácilmente convertirse en) una explicación completa de la caída a tierra del satélite, pero que era sólo una explicación parcial de la caída del satélite en Australia Central. Sin duda el acontecimiento concreto de la caída del satélite en tierra era fácticamente el mismo acontecimiento concreto de la caída del satélite en Australia Central [5]. Pero según la

[5] Debe observarse que Hempel no está seguro de que esté del todo clara la noción de "acontecimiento concreto" [10].

explicación de Hempel lo que se explica no es un acontecimiento concreto sino un hecho. Un hecho es lo que es expresado mediante una sentencia verdadera, y en tanto que las sentencias pueden ser más o menos específicas podemos hablar de hechos que son más o menos específicos.

Una de las razones por las que se producen choques entre Hempel y los críticos del modelo de explicación de la cobertura legal, tales como Scriven, es que los críticos a menudo denominan "explicación" a algo que Hempel denominaría "explicación parcial". Es evidente que el lenguaje ordinario apoya ambas posiciones. En nuestro ejemplo relativo a una parte del satélite de la Tierra que caía en Australia Central, sería del todo correcto decir que el argumento era una "explicación" del hecho de que esa parte *cayese* en Australia Central, pero igualmente se podía argüir que no era explicación en absoluto, ya que no explicaba por qué la parte en cuestión del satélite caía en *Australia Central.* Por tanto Hempel tiene perfecto derecho a utilizar la palabra "explicación" de un modo más preciso, de modo que el ejemplo mencionado no constituya una explicación sino solamente una explicación parcial del hecho en cuestión. Las críticas al modelo de explicación de la cobertura legal, en el sentido más preciso de Hempel, no tienen sentido si simplemente afirman que las explicaciones *en el sentido más amplio* de la palabra no siempre se ajustan al modelo de Hempel. Para que lo tuvieran sería necesario demostrar que Hempel, al intentar hacer más preciso el concepto de explicación, lo que hizo fue desvirtuarlo. (Discutiremos algunas objeciones de este tipo posteriormente en este capítulo.)

El ejemplo del satélite de la Tierra fue un ejemplo de explicación deductiva-nomológica parcial. Hay dos motivos por los que será útil proporcionar ahora un ejemplo de una explicación inductivo-estadística. En primer lugar las explicaciones estadísticas parciales son comunes en las partes más históricas de la ciencia, tales como la teoría de la evolución. En segundo lugar, es importante comprobar que no hace falta que construyamos la explicación parcial de tal modo que todas las explicaciones parciales sean consideradas como únicamente parciales. Hempel trata de construir "la explicación parcial" de modo que las explicaciones estadísticas puedan ser parciales o completas según el caso. El se basa en que la física moderna hace que las leyes estadísticas sean las definitivas, y, por tanto, una explicación física ya no puede ser considerada siempre como de "segunda categoría", que se pueda sustituir quizá algún día por una explicación deductivo-nomológica. Supóngase que las leyes de la mecánica cuántica nos permiten predecir que el 80 por 100 de los electrones que pasan por un cristal se adhieren a una pantalla en determinada región R de la pantalla. Entonces, si un solo electrón se adhiere a la región R, las leyes acabadas de mencionar nos permiten proporcionar una explicación estadística completa de la adherencia del electrón a la pantalla en la región R. Proporcionará, sin embargo, una explicación estadística solamente parcial del hecho de que un hombre pelirrojo observe cómo el electrón se adhiere a la pantalla en R. Así pues, la distinción entre explicacio-

nes completas y parciales es aplicable a las explicaciones inductivo-estadísticas al igual que a las deductivo-nomológicas.

Esbozos de explicación

Hempel ha introducido también el valioso concepto de esquema de explicación. A menudo se realiza la explicación de forma elíptica: es decir, algunas de las premisas no se establecen explícitamente quizás, pero se desprenderán claramente del contexto. A veces, sin embargo, no se verá claro cuáles son las premisas nuevas que harían falta para completar el argumento intentado dando lugar a una explicación completa o parcial debidamente formulada. En este caso, el argumento es más bien la sugerencia de una explicación, y Hempel denomina a tal tipo de argumento esbozo de explicación. Por supuesto, el que un trozo de discurso se interprete como una explicación completa formulada de forma elíptica, o quizá como una explicación parcial, o también quizá como simplemente un esbozo de explicación es cuestión de opinión: los diversos conceptos no son fáciles de deslindar, ya que a menudo podemos elegir libremente entre considerar que una premisa se sobreentiende tácitamente o el considerar que no es premisa en absoluto. Sin embargo, el siguiente caso nos proporcionará un claro ejemplo de algo que podría denominarse esbozo de explicación. Supóngase que se pregunta por qué una persona tiene los ojos azules y que se indica que ello se debe a sus genes. Quizá se considera a los genes simplemente como "algo" que en circunstancias ambientales normales da lugar a diversas características fenotípicas (por ejemplo los ojos azules). Es evidente que la respuesta a la pregunta "¿por qué tiene él ojos azules?" no alude a ningún tipo de regularidad genética. Lo que se nos da es simplemente un esbozo de explicación: una sugerencia para buscar una explicación dentro del campo de la genética. Es evidente que alguien que considerase "se debe a sus genes" como explicación encontraría aquí una excepción al modelo legal de cobertura, pero el proponente del modelo legal de cobertura tiene respuesta a ello, a saber, clasificando la explicación putativa como simplemente un esbozo de explicación.

La explicación como analogía

Los escritores que pudieran admitir que la explicación se ajusta formalmente al modelo legal de cobertura quizá mantengan, sin embargo, que existen requisitos adicionales, de tal modo que las leyes explicativas deben guardar *analogía* con leyes ya aceptadas. El físico N. R. Campbell, que escribió algo muy interesante sobre la filosofía de la física hace una generación aproximadamente, insistió en este punto [11]. Considérese la explicación de la ley de Boyle en términos de la teoría cinética de los gases, en la cual un gas se considera constituido por muchísimas partículas per-

fectamente elásticas que se mueven de acuerdo con las leyes de la mecánica newtoniana. En tanto en cuanto la ley de Boyle es deducible de estas concepciones mecánicas se ajusta al modelo deductivo-nomológico, pero Campbell argumentaba que a menos que este modelo formal de explicación estuviese supeditado al requisito adicional de la analogía, tendrían que considerarse explicaciones diversos argumentos del todo triviales. Así, utilizando su ejemplo [12], la constancia del radio de la resistencia eléctrica de un trozo de metal respecto a su temperatura absoluta podría ser "explicada" mediante la siguiente teoría:

$$\text{Resistencia } (R)=(c^2+d^2)\,a$$

$$\text{Temperatura } (T)=\frac{cd}{b}$$

a y b son constantes y $c=d$

De esta teoría deducimos que

$$\frac{R}{T}=(c^2+d^2)\,a\left(\frac{b}{cd}\right)=2ab=\text{constante}$$

Esta teoría es evidentemente trivial, no proporcionando ninguna explicación real de la constancia de $\frac{R}{T}$. Lo que la hace trivial, según Campbell, es su falta de analogía con lo que ya conocemos. Contrástese con la teoría cinética de los gases en la que las partículas postuladas se comportan exactamente igual que *si* fuesen pequeñas bolas de billar perfectamente elásticas.

Se necesitan requisitos adicionales (además y por encima de la capacidad de proporcionar deducciones de los "explicanda") indudablemente, antes de que una teoría pueda ser explicativa y no meramente trivial. No era necesario, sin embargo, que Campbell estableciese el requisito de la *analogía* para lograr esto. Las teorías que consideramos auténticamente explicativas pueden utilizarse para predecir hechos distintos a aquéllos para los que fueron en principio inventadas, o pueden modificarse de un modo sencillo a fin de que lo hagan. Es decir, la diferencia importante entre la teoría trivial del ejemplo de Campbell y la teoría cinética de los gases, que no es trivial, puede decirse que radica simplemente en la capacidad predictiva de esta última o, al menos, que está íntimamente relacionada con ella, y no que tenga nada que ver en especial con la analogía. Por supuesto, cuando Campbell discute un tipo más fenomenológico de teoría, ejemplificado por la teoría de Fourier de la conducción del calor, sustituye el requisito de la analogía por el de la simplicidad [13].

No puede negarse que la analogía ha sido, con frecuencia, de gran valor heurístico para la construcción de teorías científicas. En algunos casos podemos ver el porqué de ello. Porque si el gas es realmente un

montón de partículas que se comportan de modo bastante semejante al de diminutas bolas de billar perfectamente elásticas, entonces no es difícil ver por qué un teórico al que se le ocurriese considerar a un gas siguiente la *analogía* de una legión de bolas de billar diminutas perfectamente elásticas tendría éxito. Más aún, la analogía puede verse en las ciencias en casos mucho menos directos. La teoría de Bohr acerca del átomo tuvo éxito, y en ella los átomos se describían siguiendo la analogía de un sistema de planetas girando alrededor de un sol. No obstante esta teoría tuvo que modificar la analogía considerablemente, ya que tuvo que imponer las condiciones cuánticas que impiden que los electrones planetarios de un átomo germinen en el interior del núcleo y que explican cosas tales como los espectros de línea discontinua. Son precisamente estas condiciones cuánticas las que han hecho que la teoría tenga tanto éxito, y si volvemos la vista atrás resultará evidente que son las partes no análogas de la teoría las que más han contribuido a su éxito y las que han superado la prueba del tiempo. Es cierto que en el caso de las teorías más elaboradas que han sustituido sucesivamente a la teoría de Bohr continúa habiendo *cierta* analogía con la mecánica clásica. Así, en el moderno desarrollo de la mecánica cuántica, a través de la abstrusa teoría matemática de los espacios de Hilbert, los físicos tienen que basarse a menudo en analogías, ya que es necesario postular ecuaciones diferenciales con relación a operadores en el espacio de Hilbert, y al calcularlas los físicos se han guiado por ecuaciones diferenciales análogas (que relacionan no operadores sino cantidades físicas) de la física clásica. Debe recordarse, sin embargo, que ésta es una cuestión heurística: que una ecuación diferencial sea o no correcta viene determinado no por su analogía con las ecuaciones clásicas sino que depende de si son o no correctas las consecuencias observables que puedan derivarse de ella. En tanto en cuanto la analogía ha resultado una buena guía puede considerarse un hecho afortunado relativo al mundo, pero también pudiera ser reflejo del hecho de que cuando los investigadores cuentan con poco más que la analogía para guiarse, entonces sus teorías inevitablemente tenderán a mostrar alguna analogía con teorías precedentes. Más aún, si en el curso del desarrollo de la física una teoría T_0 es sustituida sucesivamente por las teorías T_1, T_2, ..., T_n, entonces fácilmente puede darse una analogía entre T_r y T_{r+1} para cada r para $0 \leqslant r < n$. Es fácil que se dé una confusión en este caso. Aun si (ya que es algo dudoso) es un hecho de la psicología humana el que los investigadores se guíen siempre por la analogía, esto no muestra que la analogía tenga nada que ver con que una teoría sea buena. Supóngase que, por casualidad o debido a la intervención de un super-genio, la mecánica cuántica se hubiera originado repentinamente en su forma más moderna, con espacios de Hilbert y todo. Entonces la cadena T_0, T_1, ..., T_n habría sido sustituida por la simple transición de T_0 a T_n. Con toda seguridad T_n no sería por ello peor o menos explicativa. Probablemente no habría sido aceptada por los científicos en general, ya que habría sido demasiado avanzada para su época, exactamen-

te lo mismo que ocurrió con la mecánica estadística de Gibbs cuando fue propugnada por primera vez. Pero a menos que adoptemos una posición pragmática acerca de la explicación, debemos decir que esta cuestión hace relación a la psicología y sociología respecto a los científicos, y no implica la bondad o maldad de las explicaciones en cuanto tales.

La explicación como reducción a lo conocido

La exigencia de analogía en las teorías quizá se derive de una concepción de la explicación como reducción a lo conocido. De este modo, si el sistema solar nos es más conocido que lo son los átomos, entonces la teoría de los átomos de Bohr quizá capte nuestra atención porque en cierto modo reduce el comportamiento de los átomos al comportamiento más conocido del sistema solar. Sin embargo, incluso la teoría de Bohr introduce ideas (tal como la del salto cuántico) que no pueden ser reducidas a nada de lo que constituya nuestra experiencia anterior. Los científicos han superado, con mucho, los ideales de W. Thomson (Lord Kelvin) del siglo pasado. Kent llegó a decir que nunca podría estar seguro de que entendía algo hasta que no pudiera hacer un modelo mecánico de ello. De este modo quiso explicar las propiedades eléctricas de la materia basándose en analogías mecánicas, mientras que, por supuesto, ahora intentamos explicar algunas de las características mecánicas de la materia (tales como la elasticidad) basándonos en las propiedades eléctricas de las partículas a partir de las cuales se constituye la materia. Ciertamente no es ideal de la ciencia moderna el que la explicación deba ser una reducción a lo conocido, puesto que las leyes básicas de la mecánica cuántica y de la teoría de la relatividad resultan muy extrañas a la intuición vulgar. Con todo, la objeción más decisiva a la consideración de la explicación como una reducción a lo conocido es la de que frecuentemente es necesario explicar hechos muy conocidos. Hempel, por ejemplo [14], menciona el caso de la paradoja de Olbers, que sugiere, basándose en determinados presupuestos plausibles, que el firmamento no debería estar oscuro por las noches, sino que, por el contrario, habría de estar realmente muy iluminado. Discutiremos la paradoja de Olbers en el capítulo 8, donde veremos los conocimientos cosmológicos que podemos lograr si reflexionamos acerca del conocido hecho de la oscuridad del firmamento nocturno y las posibles explicaciones de ello. Existen, por descontado, todo tipo de hechos conocidos (así como desconocidos) acerca de los cuales los científicos buscan una explicación; por ejemplo, por qué los cuerpos caen, por qué el hierro se enmohece, por qué necesitamos sal en nuestra dieta y por qué los hijos se parecen a sus padres. Las leyes y generalizaciones que se utilizan para explicar tales hechos desde luego se convierten en algo finalmente familiar para los científicos que las utilizan, pero si la explicación fuese la reducción a lo conocido, es difícil explicar por qué habríamos de necesitar jamás la explicación de los hechos conocidos.

La explicación como redescripción

Considérese el caso del enmohecimiento del hierro, mencionado anteriormente. Parte de la explicación del enmohecimiento consiste en la identificación del enmohecimiento y la oxidación. Entonces la explicación de un caso de enmohecimiento requerirá el hecho de que el vapor de agua esté presente junto con una ley que determine que cuando esté presente el vapor de agua tendrá lugar la oxidación del hierro. Podríamos interpretar la situación mencionada de dos formas: 1) El científico cambia el significado de "enmohecimiento" en su acepción vulgar de "enrojecimiento del hierro expuesto a la humedad" por el de "oxidación". Y entonces el que explica es el de la oxidación del hierro. 2) Lo que es explicado es el enrojecimiento del hierro, con la ayuda de una ley adicional, "todos los casos de enrojecimiento del hierro por exposición a la humedad son casos de oxidación".

Veamos qué diferencias se dan, si se da alguna, entre las interpretaciones 1) y 2) acabadas de indicar. Según 1) el "quid" de la explicación radica en el cambio de significado de la palabra "enmohecerse", o, en otras palabras, en la *redescripción* del enrojecimiento producido por la exposición a la humedad o enmohecimiento. Según 2), el "quid" radica no en la redescripción, sino en la ley adicional que se necesita. Mas, ¿no podría mantenerse, sin embargo, que aceptar la *ley adicional* de 2) no es ni más ni menos que aceptar la *redescripción*? La consideración de la explicación como redescripción no parece diferir del modelo de cobertura legal.

Por otra parte, aunque quizá no se dé inconsistencia formal entre un filósofo que insista en el modelo deductivo y otro que insista en la redescripción, es bien posible que aparezca una gran diferencia externa. El interés por el modelo deductivo va, naturalmente, acompañado de (aunque no tiene por qué ser así *necesariamente*) una filosofía de la ciencia instrumentalista, filosofía que valora a la ciencia principalmente como instrumento para predecir nuestras experiencias, mientras que la consideración de la explicación en términos de redescripción puede ir acompañada de una concepción más desinteresada de la ciencia, la concepción de la ciencia como ayuda para comprender cómo es el mundo, por el solo gusto de lograr esta comprensión, y completamente al margen de cualquier tipo de valor predictivo o práctico que la ciencia pueda ofrecernos [15].

La crítica de Feyerabend al modelo deductivo

Las consideraciones acabadas de mencionar sugieren un punto de vista acerca de la explicación que mantiene que lo que es importante en la explicación no es tanto su forma (que es explicada por el modelo deduc-

tivo) sino la idea de que una explicación proporciona comprensión en tanto en cuanto nos ofrece una descripción más satisfactoria de los hechos. Por supuesto que si "más satisfactoria" significa simplemente que "gusta más", entonces esta concepción de la ciencia tendría poco valor: lo que le gusta más a un curandero no es lo que le gusta más a un médico; por lo que debemos buscar una explicación de "mas satisfactoria" que sea más objetiva. Valdrá la pena que consideremos algunos puntos de vista interesantes que han sido propuestos por Paul Feyerabend [16].

Consideremos las leyes de la termodinámica fenomenológica (o "clásica"). Estas explican algunos hechos conocidos y otros con los que no estamos familiarizados. Por ejemplo, el que el hielo se derrita en una copa de licor es explicado por la segunda ley, que afirma que en un sistema cerrado la entropía aumenta siempre, junto con determinados presupuestos fácticos, tales como que la temperatura de la habitación en la que se encuentre la copa de licor esté por encima del punto de congelación, y que el sistema que contenga la copa de licor, junto con determinado volumen del aire circundante, pueden ser considerados como un sistema cerrado aproximadamente. Ahora podríamos considerar la cuestión de si las propias leyes de la termodinámica fenomenológica pueden ser explicadas. De hecho se dice a menudo que la teoría de la termodinámica fenomenológica puede ser "reducida" a la teoría estadística, considerándose generalmente a la explicación o "reducción" de acuerdo con un modelo deductivo. Es decir, una teoría T se reduce a, o se explica mediante, una teoría T' si T es deducible de T', junto con determinadas *reglas de correspondencia.* Estas reglas de correspondencia son necesarias debido a que por lo menos algunos de los conceptos de T y T' son diferentes; así, el concepto de temperatura en la termodinámica clásica se iguala al del promedio de la energía cinética de las moléculas, que es un concepto de la termodinámica estadística. Además, generalmente se sostiene que exactamente al igual que la termodinámica estadística tiene que vincularse a la termodinámica clásica mediante reglas de correspondencia, así también son necesarias las reglas de correspondencia para vincular los términos de la termodinámica clásica a los del uso vulgar. Las observaciones mediante las cuales comprobamos las teorías deberían establecerse, según esto, en lenguaje puramente vulgar: no como "la temperatura de este líquido es 30° C" sino como "el líquido contenido en este tubo estrecho de vidrio llega hasta una raya marcada como '30'". Contamos ahora con una imagen jerárquica de la explicación: 1) Las observaciones vulgares se explican mediante la deducción a partir de la termodinámica estadística junto con determinadas reglas de correspondencia. 2) La termodinámica clásica se explica mediante la deducción a partir de la termodinámica estadística junto con determinadas reglas de correspondencia.

La jerarquía acabada de mencionar consta de tres elementos, pero no hay razón que impida que existan jerarquías de cuatro, cinco o más elementos. Sin embargo, la de tres elementos ha logrado un puesto destacado en el moderno pensamiento científico debido a una influyente interpreta-

ción de la mecánica cuántica según la cual todas las explicaciones de la física atómica deben adoptar la forma: mecánica cuántica → física clásica → observaciones vulgares. Los físicos que aceptan esta interpretación mantienen que es imposible que la mecánica cuántica pueda ser comprobada directamente mediante observaciones, ya que mantienen que al menos parte de la teoría del aparato experimental debe ser clásica. Otra muestra importante de un orden jerárquico en la deducción es: relatividad general → relatividad especial → mecánica newtoniana → observaciones de los movimientos planetarios.

Feyerabend rechaza este modelo de explicación: 1) Declara que cuando se dice que una teoría T se "explica" mediante una teoría T', lo que realmente ocurre es que T no ha sido explicada en absoluto, sino que ha sido *sustituida* por T'. (Lejos de haber sido *explicada* por la teoría T', lo que se ha demostrado es que la teoría original T era falsa.) 2) Mantiene que pueden deducirse consecuencias observacionales de T y de T' *directamente* (en sus palabras "una teoría es su propio lenguaje observacional") y, por tanto, que las reglas de correspondencia son innecesarias.

La primera crítica de Feyerabend a la concepción jerárquica es la de que cuando una teoría (tal como la termodinámica fenomenológica) se considera reducida a otra (tal como la termodinámica estadística) las dos teorías son frecuentemente de hecho inconsistentes entre sí. Suponiendo que las dos teorías sean internamente consistentes debemos concluir que una no puede ser *deducida* de la otra.

Tomado esto aisladamente, a primera vista no parece ser un ataque muy perjudicial al modelo deductivo. J. Kemeny y P. Oppenheim, que se aproximan mucho más a Hempel que a Feyerabend en su filosofía de la ciencia, han mantenido algo semejante [17]. Hempel [18] está completamente de acuerdo en que a veces un hecho idéntico podría ser subsumido en teorías mutuamente incompatibles. Más aún, un filósofo como Nagel [19], que da una interpretación deductiva de la explicación de las leyes (o teorías) podría replicar que no son las propias leyes lo que se deduce sino su *aproximada corrección*. De este modo la relatividad general nos permite deducir la corrección aproximada de la ley de la gravitación de Newton, y la teoría ondulatoria de la luz nos permite deducir la corrección aproximada de la óptica geométrica. (Hilary Putnam, que defiende a Nagel en este sentido, explica "aproximadamente verdadera" del modo siguiente: Decir que una teoría es aproximadamente verdadera es decir que aunque la propia teoría no sea verdadera, "determinada *consecuencia lógica* de la teoría obtenida, por ejemplo, mediante la sustitución de "iguales" por "iguales más o menos delta" es verdadera" [20].)

¿Es correcto decir, sin embargo, que la teoría de la gravitación de Newton o la de la óptica geométrica es aproximadamente verdadera siquiera? La ley del cuadrado inverso de Newton no es ni siquiera aproximadamente verdadera cerca de un cuerpo muy pesado y denso. La relatividad general predeciría una gran desviación de la ley del cuadrado inverso cerca de un "quasar". Asimismo, la óptica geométrica proporciona

resultados tremendamente incorrectos respecto a lo que ocurre cuando la luz es refractada a través de una o más ranuras. Lo aproximadamente verdadero no es la ley original de Newton sino *otra* ley que podemos obtener a partir de la ley original mediante la inserción de la frase "lejos de cualquier cuerpo pesado y denso". Pero esto equivale a admitir que lo que se deduce a partir de la relatividad general no es la teoría newtoniana sino el hecho de que la teoría newtoniana proporciona predicciones aproximadamente correctas en determinadas circunstancias (el porqué de su funcionamiento en la práctica así como también el cómo de su funcionamiento). Esto es, en parte, la posición del propio Feyerabend. Sin embargo, es difícil de creer que el propulsor de la consideración jerárquica de la ciencia no admitiese estas modificaciones de su propio punto de vista.

La objeción más fuerte de Feyerabend [6] parece radicar más bien en su afirmación de que las observaciones vulgares no pueden ser formuladas en un lenguaje teórico neutral, que una teoría es su propio lenguaje observacional. Evidentemente, no va desencaminado. Es cierto que las observaciones vulgares pueden estar cargadas de teoría, es decir, pueden utilizar una terminología que es explicable sólo dentro del esquema conceptual de alguna teoría. Contrariamente a lo que sostenía Eddington [22], los enunciados observacionales *no* son enunciados acerca de una determinada interpretación. Un campesino sin preparación científica podría indicar, ciertamente, que un objeto negro semejante a una aguja señalaba a la cifra "35" en un objeto redondo semejante a un reloj. Lo que *no* podría hacer es indicar que la corriente a través del miliámmetro era de treinta y cinco miliamperios, ya que no contaría con el concepto de corriente eléctrica, y mucho menos con el de amperio. Ni tampoco podría servir la declaración de un campesino como prueba de una hipótesis científica [7]. El miliámmetro podría no estar correctamente calibrado, o podría tener una resistencia paralela a él, o podría estar incluso vacío y tener dentro de él un artefacto que hiciese maniobrar el marcador. La declaración del campesino sería compatible con todas estas posibles situaciones. El escrutinio de una serie de observaciones científicas nos mostraría la carga teórica que contienen incluso las observaciones más simples.

Nos encontramos ahora en la situación aparentemente paradójica de que si una sentencia observacional O ha de servir como prueba determinante de dos teorías conceptualmente distintas, T_1 y T_2, entonces O debe expresar en realidad dos proposiciones diferentes, O_1 y O_2. Es decir, al expresar O_1 está cargada con los conceptos de la teoría T_1, y al expresar O_2 está cargada con los conceptos de la teoría T_2.

---

[6] Y de la posición de Wilfrid Sellars también, que es algo intermedio entre la de Nagel y la de Feyerabend [21].

[7] Sin embargo, como Feyerabend señala [23], nosotros, si tuviésemos conocimientos acerca de la electricidad y los amperios, podríamos *utilizar el hecho de que* el campesino haya hecho tal declaración como comprobación de una hipótesis eléctrica.

Ahora bien, la imagen habitual de un experimento crucial es como sigue (siendo T_1 y T_2 las teorías rivales, A la conjunción de las proposiciones que expresan condiciones limítrofes y O una sentencia observacional):

$$(A \cdot T_1) \supset O$$
$$(A \cdot T_2) \supset \sim O$$
$$O \cdot A$$
$$\therefore \sim T_2$$

Según la opinión de Feyerabend esto ha de ser sustituido por el modelo:

$$(A \cdot T_1) \supset O_1$$
$$(A \cdot T_2) \supset \sim O_2$$
$$O_2 \cdot A$$
$$\therefore \sim T_2$$

Obsérvese que O_1 y O_2 serán morfológicamente la misma sentencia, pero consideramos a O tal como funciona en el contexto de T_1 como una sentencia O_1 y a O tal como funciona en el contexto de T_2 como una sentencia diferente O_2.

(Quizá A debiera bifurcarse en dos sentencias, a saber A_1 y A_2, al igual que O se bifurca en O_1 y O_2, pero podemos prescindir de esta matización.)

Ahora bien, para refutar la teoría T_2 el defensor de T_1 debe afirmar O_2. (Podemos presuponer que la verdad de A no es cuestionable.) Pero lo que el defensor de T_1 estará afirmando no será O_2 si no O_1. Feyerabend señala, sin embargo, que O_1 y O_2 son pragmáticamente equivalentes [8] en el sentido de que exactamente la misma situación que hace que un defensor de T_1 afirme O_1 hará que un defensor de T_2 afirme O_2.

De este modo puede Feyerabend explicar experimentos cruciales entre teorías incompatibles y conceptualmente distintas que es congruente con su punto de vista de que una sentencia observacional tiene significados distintos en el contexto de cada una de las teorías. Lo que el defensor de T_1 hace es considerar de modo provisional lo que el defensor de T_2 tendría que decir si hubiera de negar la proposición O_2, que es el equivalente pragmático de O_1.

A veces una teoría T_1 logra subsistir después de una comprobación que no es ni siquiera posible dentro del contexto de una teoría menos satisfactoria, T_2. Un buen ejemplo utilizado por Feyerabend es el del movimiento browniano [25]. Un ejemplo del movimiento browniano es el movimiento errático que se puede observar en granos de polen microscó-

[8] Esta explicación pragmática de los enunciados observacionales es semejante a la formulada en principio por Carnap [24], pero al parecer ha sido desechada últimamente por éste.

picos suspensos en un líquido[9]. Este fenómeno es una violación de la termodinámica clásica, ya que los granos de polen, junto con el líquido circundante, constituyen un mecanismo de calentamiento de un tipo que no es admisible de acuerdo con la segunda ley de la termodinámica clásica. Sin embargo, no es posible una comprobación *experimental* directa de la termodinámica fenomenológica con el fenómeno del movimiento browniano; sería del todo imposible crear termómetros que fuesen lo bastante sensibles para medir los intercambios térmicos entre el polen y el fluido o que no produjesen ellos mismos intercambios térmicos que alterarían las cantidades que habían de ser medidas. Sin embargo, desde el punto de vista de la termodinámica estadística puede demostrarse que el movimiento browniano implica violaciones de la segunda ley de la termodinámica (fluctuaciones que implican un *descenso* temporal de la entropía de un sistema).

Este ejemplo de Feyerabend muestra que la validez de una teoría no viene determinada únicamente por predecir aproximadamente los resultados observacionales de alguna teoría anterior de aceptación general, sino que lo que la hace válida es que esté confeccionada de modo que compruebe *directamente* hechos observacionales, dando lugar incluso a que la teoría anterior pueda ser rechazada tomando como base alguna comprobación observacional que sería del todo imposible en tanto que pensáramos dentro del esquema conceptual de la vieja teoría. Por supuesto, el hecho de que podamos deducir, a partir de la nueva teoría, que la teoría antigua sería válida dentro del dominio observacional en el que de hecho *ha* sido válida, nos permite considerar estas comprobaciones observacionales de la antigua teoría como comprobaciones observacionales válidas *ipso facto* de la nueva teoría.

Por ejemplo, como el propio Feyerabend concede, sería imposible, *en la práctica,* utilizar la relatividad general en la mecánica celeste (es decir, la teoría de los movimientos detallados de los planetas y sus satélites alrededor del Sol). Ha de utilizarse la mecánica newtoniana, y la relatividad general aparece únicamente en aquellos casos (tales como el avance del perihelio de Mercurio) en los que parece ser que la teoría newtoniana proporciona predicciones incorrectas. Pero en tanto que pueda deducirse de la relatividad general que la mecánica newtoniana proporciona predicciones aproximadamente exactas en aquellas circunstancias en que la mecánica newtoniana *ha* sido válida, la validez de la mecánica newtoniana puede ser considerada como validez también de la teoría general de la relatividad.

---

[9] Los fenómenos cuantitativos del movimiento browniano fueron predichos correctamente por Einstein, basándose en los cálculos teóricos de la mecánica estadística y la teoría cinética de la materia, que son los principales pilares de la termodinámica estadística [26].

Crítica del uso, por parte de Feyerabend, del concepto de "significado" (meaning)

Muchos semánticos y filósofos modernos se muestran recelosos respecto a la palabra "significado" (meaning): lo que queremos manifestar mediante ella no está muy claro, y Quine incluso postula que prescindamos de "significados", el "mismo significado" y "distinto significado" por completo.

Considérese el concepto de temperatura. De acuerdo con la termodinámica estadística ésta consiste en determinada propiedad estadística. Por ejemplo, la temperatura de un gas es la energía cinética media de sus moléculas. En la termodinámica clásica la temperatura se define de un modo completamente distinto, a saber, en términos de intercambios térmicos en la operación de un perfecto mecanismo térmico. Con anterioridad a la termodinámica clásica la temperatura era simplemente lo que quiera que fuese lo que los termómetros medían. Parece, por consiguiente, que está justificado, en parte, el alegato de Feyerabend de que las observaciones realizadas en los contextos de teorías diferentes (por ejemplo, registros de temperatura tales como los realizados respectivamente por Galileo y por los defensores de la termodinámica clásica y estadística) poseen distinto significado. Hilary Putnam ha llamado, desde luego, la atención acerca de un mal razonamiento de Feyerabend al decir que Galileo expresaba algo distinto mediante el término "temperatura" de lo que expresamos ahora con este término, a saber, que el termómetro mide no la temperatura sino una función tanto de la temperatura como de los efectos de la presión atmosférica, las diferencias en el comportamiento de los fluidos de diversas composiciones químicas que podrían utilizarse en el termómetro y demás [27]. Según Putnam, mediante "temperatura" Galileo expresaba simplemente alguna propiedad intrínseca de una sustancia que el termómetro mide, no necesariamente con exactitud, y que Galileo comprendería tan bien como nosotros el enunciado de que la temperatura medida no corresponde exactamente a la temperatura real [28]. Entonces probablemente, de acuerdo con Putnam, el defensor de la termodinámica clásica y el defensor de la termodinámica estadística, expresarán mediante "temperatura" lo mismo que expresaba Galileo, a saber, la propiedad intrínseca de los cuerpos que los termómetros miden, no necesariamente con exactitud, y una "definición" de la temperatura dada por los físicos teóricos es, en realidad, un enunciado fáctico que afirma que la propiedad dada en la definición es de hecho la propiedad a la que se alude en la definición de Galileo. (Esto se asemeja mucho a la afirmación fáctica de que la propiedad de ser rojo es la propiedad común a exactamente tres coches que estén fuera en la calle [29].) No veo, sin embargo, como esta alegación pueda ser apoyada por las explicaciones acerca de la temperatura que aparecen en los tratados científicos [10].

[10] Resulta instructivo al respecto mirar la definición de *temperatura* en *The*

Los conceptos teóricamente neutrales y el lenguaje vulgar

En la sección precedente, Putnam, como pudimos ver, proporcionaba una explicación del concepto de temperatura que era neutral en relación a diversas teorías. Si ésta era una explicación correcta entonces el concepto de temperatura podría denominarse teóricamente neutral. Mantener que ciertos conceptos poseen esta característica es una posición que Feyerabend ha denominado "teoría hueca del significado" [31]. Yo considero que el nombre es inapropiado, ya que no se trata de una teoría del significado sino de la explicación del significado de determinados conceptos. Las objeciones de Feyerabend son de tipo metadológico. Cuanto más ricos sean nuestros conceptos más fáciles de comprobar serán nuestras afirmaciones. Esta es, sin duda, una buena razón del porqué en ciencias tratamos de que muchas de nuestras sentencias *no* sean teóricamente neutrales, pero no demuestra que no sea alguna vez conveniente que *algunas* de ellas sean teóricamente neutrales. Supongamos que la prueba decisiva para optar entre la relatividad general y la mecánica newtoniana consista en que si la primera es correcta entonces nuestro telescopio, cuando realicemos determinada observación, apuntará hacia una chimenea determinada y no hacia una palmera determinada. (Este supuesto es a duras penas realístico, considerando la minuciosidad y cuidado que de hecho requeriría la observación, pero permítaseme partir de él a fin de hacer más clara esta explicación.) ¿Significa esto que "chimenea" significa algo distinto en el contexto de cada una de estas teorías? Quizá sea así si en la explicación einsteiniana una chimena es una invariante cuatri-dimensional en el espacio-tiempo, mientras que en la explicación newtoniana es un sólido espacio tri-dimensional que permanece a través del tiempo [11]. Mas, ¿por qué no considerar la "chimenea" significando algo neutral con relación a estas teorías, definida únicamente en términos de la función arquitectónica a la que están destinadas las chimeneas? Y de igual modo se podría dar una definición botánica de "palmera" que fuese neutral respecto a las teorías físicas [32].

Tampoco tenemos por qué seguir a Feyerabend cuando afirma que el lenguaje vulgar implica falsedades. A veces sí las implica, con toda seguridad, como cuando decimos "lleva el criket en las venas: su padre fue un jugador de criket muy bueno y su abuelo materno capitaneó al Australia". Ahí realmente deberíamos decir "en sus genes" y no "en sus venas". (¿Resultan pedantes estas disquisiciones? No lo estimo así, en tanto que los sudafricanos blancos se opongan a que se les hagan transfusiones de sangre con sangre de negros.) Con todo, es erróneo mantener

---

Penguin Dictionary of Science [30], que está dada en términos de la proporción de variaciones caloríficas, que nos lleva a la definición de *calor*, que está dada en términos de *energía*. Con lo cual entramos inmediatamente en lo teórico.

[11] No está del todo claro, sin embargo, que esto suponga en realidad una diferencia, ya que no hay nada que impida al propio newtoniano el pensar en términos de espacio-tiempo.

que el lenguaje vulgar, o cualquier otro lenguaje, es falso, ya que es de suponer que la contradictoria de una sentencia esté en el lenguaje en que está formulada, y o bien esa sentencia o su contradictoria han de ser verdaderas [33]. Podemos incluso utilizar el *lenguaje* de la teoría flogística para afirmar verdades, como por ejemplo "el flogisto no existe" *.

Si gran parte del lenguaje vulgar es teóricamente neutral entonces es posible formular las observaciones de un modo teóricamente neutral. Más aún, incluso en los casos en que las sentencias observacionales no sean teóricamente neutrales, el modelo de comprobación de la teoría simplemente se bifurcará en *dos* modelos deductivos. (Al igual que en la página 67.) Igualmente el modelo formal de explicación seguirá también el modelo deductivo hempeliano. En realidad no se da auténtica contradicción entre la interpretación de la explicación de Hempel en términos de su modelo deductivo y la insistencia de Feyerabend en la explicación como redescripción. Podemos redescribir el enmohecimiento como oxidación, por ejemplo, y luego deducir un caso particular de oxidación de la teoría química junto con determinadas condiciones iniciales. Hempel será partidario de complicar esta interpretación manteniendo algún significado vulgar o teóricamente neutral de "enmohecerse" y ampliando la explicación con la introducción de la proposición "Todos los casos (¿o la mayoría de ellos?) de enmohecimiento (en el sentido vulgar) son casos de enmohecimiento en el sentido de oxidación". Lo cual no alterará el hecho de que tenga lugar la explicación cuando el científico vea un objeto enmohecido como un caso de oxidación.

Decididamente, lejos de considerar la interpretación de la explicación como redescripción distinta a la interpretación de Hempel, podemos decir que la redescripción científica está esencialmente relacionada con la deducción a partir de leyes (que es en lo que Hempel insiste). El ideal de la ciencia no es la descripción del mundo en el sentido de un inventario general, sino el de una descripción *simple,* que significa que toda suerte de fenómenos deben poder deducirse de sólo unas cuantas leyes.

No creo, por tanto, que se dé inconsistencia formal entre la teoría de la explicación de Hempel y la de Feyerabend. A Feyerabend, sin embargo, le preocupan especialmente no tanto las explicaciones rutinarias dentro de teorías bien experimentadas, sino aquellos casos de *progreso* en la explicación, cuando abandonamos alguna teoría en favor de otra mejor [12].

---

* "Flogisto": principio imaginado por Sthal, médico y químico alemán muerto en 1734, para explicar los fenómenos caloríficos y que él suponía que se desprendía de los cuerpos durante la combustión. *(Nota de la Traductora.)*

[12] Me he ocupado de aquellas partes de la teoría de la explicación de Feyerabend que parecen colisionar con la teoría de Hempel, y he intentado conciliar ambas. Sin embargo, los escritos de filósofos tales como Hempel o Nagel, por una parte, y Feyerabend, por la otra, presentan un matiz muy distinto: los de los primeros sugieren una actitud más bien conservadora en relación con las teorías científicas, mientras que los de este último suponen una posición intensamente inconformista. Tampoco he hecho justicia a Feyerabend por lo que respecta a la riqueza de sus observaciones perspicaces y a las sugestivas conjeturas que se encuentran en sus escritos.

Sir George Thomson ha planteado la cuestión de si el objetivo primordial de la ciencia es el de proporcionar explicaciones [34]. Es por lo menos tan importante, según él cree, el descubrimiento de *nuevos efectos*. De modo semejante, yo indicaría que lo más interesante de la ciencia es que nos dice lo que el mundo es mediante teorías simples y unificadas. Por tanto, si se pueden describir científicamente nuevos efectos (o viejos efectos, que también valdría para este ejemplo), esto sólo puede hacerse presentando alguna teoría sobre ellos, que será según el modelo deductivo, y que por tanto proporcionará explicaciones según el modelo hempeliano. Por tanto, decir que el objetivo primordial de la ciencia es proporcionar explicaciones y decir que su objetivo es describir el mundo no significa decir, en realidad, cosas muy diferentes.

Sugerencias respecto a lecturas posteriores

Con relación a la teoría de la explicación de C. G. Hempel, la mejor formulación, y a la vez la más reciente, es el capítulo 12 "Aspects of Scientific Explanation" de su obra *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (New York: Free Press, 1964), pero los capítulos 9-11 se ocupan también de este tema. La posición de Nagel puede verse en su *Structure of Science* (New York: Harcourt, Brace and World, 1961), especialmente los capítulos 2, 3 y 11. Yo he comparado la posición de Nagel con la de Paul Feyerabend y Wilfrid Sellars, siendo comentada mi exposición por Hilary Putnam y por Sellars, trabajos que a su vez han sido comentados por Feyerabend: véase J. J. C. Smart, "Conflicting Views about Explanation" en R. S. Cohen y M. W. Wartofsky (eds.), *Boston Studies in the Philosophy of Science*, Vol. II (New York: Humanities Press, 1965), págs. 157-169; Wilfrid Sellars, "Scientific Realism or Irenic Instrumentalism", *ibíd.*, págs. 171-204; Hilary Putnam, "How not to talk about Meaning", *ibíd.*, págs. 205-222; y Paul Feyerabend, "Reply to Criticism", *ibíd.*, págs. 223-261. También con relación a Sellars véase Wilfrid Sellars, *Science, Perception and Reality* (London: Routledge & Kegan Paul, 1963, y New York: Humanities Press, 1963, traducción española: *Ciencia, percepción y realidad*, ed. Tecnos, Madrid, 1971), capítulos 1 y 4 [el capítulo 4 apareció primeramente en H. Feigl y G. Maxwell (eds.), *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart y Winston, 1961), en donde se inserta también un comentario a cargo de Feyerabend]. En relación con Feyerabend véanse especialmente sus trabajos: "Explanation, Reduction, and Empiricism" en H. Feigl y G. Maxwell (eds.), *Scientific Explanation, Space and Time, Minnesota Studies in the Philosophy of Science*, Vol. III (Minneapolis: University of Minnesota Press, 1962), págs. 22-97; "How to be a Good Empiricist— A Plea for Tolerance in Matters Epistemological", en B. Baumrin (ed.) *Philosophy of Science, The Delaware Seminar*, Vol. 2 (New York: Interscience Publishers, 1963), págs. 3-39; y "Problems of Empiricism", en R. G. Colodny (ed.), *Beyond the Edge of Certainty* (Pittsburgh: University of Pittsburgh Press, 1965), págs. 145-260. Las consideraciones de Feyerabend con relación a la explicación son criticadas en Dudley Shapere, "Meaning and Scientific Change", en R. G. Colodny (ed.) *Mind and Cosmos* (Pittsburgh: University of Pittsburgh Press, 1966), págs. 41-85. Otros artículos relevantes son: P. Achinstein, "On the Meaning of Scientific Terms", *Journal of Philosophy*, LXI (1964), 497-509, P. K. Feyerabend, "On the Meaning of Scientific Terms", *ibíd.*, LXII (1965), 266-274, y Robert E. Butts, "Feyerabend and the Pragmatic Theory of Observation", *Philosophy of Science*, XXXIII (1966),

383-394. Entre los artículos más breves que los indicados, acerca de la explicación, se encuentran: John Hospers, "What is Explanation?" en Antony Flew (ed.), *Essays in Conceptual Analysis* (London: Macmillan, 1956, y New York: St. Martin's Press, 1956), págs. 94-119, que insiste en el modelo deductivo; y G. Schlesinger, "What is Science for?", *Australian Journal of Science,* XXVI (1963), 163-167, que considera la explicación como re-descripción. Una concepción muy poco frecuente es la sostenida en el interesante trabajo de Sylvian Bromberger, "A Theory about the Theory of Theory and about the Theory of Theories", en B. Baumrin (ed.), *The Delaware Seminar,* Vol. 2, *op. cit.,* págs. 79-106, y en su artículo "Why-Questions", en R. G. Colodny (ed.), *Mind and Cosmos, op. cit.,* págs. 86-111. Michael Scriven es uno de los principales defensores de la concepción contextualista de la explicación. Véase su trabajo "Explanations, Predictions and Laws", en H. Feigl y G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science,* Vol. III, *op. cit.,* págs. 170-230. May Brodbeck en el mismo volumen, págs. 231-272, defiende el modelo deductivo de la explicación contra las objeciones hechas anteriormente por Scriven. Otro escrito a cargo de Scriven es "The Limits of Physical Explanation", en B. Baumrin (ed.), *The Delaware Seminar,* Vol. 2, *op. cit.,* págs. 107-135. Acerca de la cuestión de distinguir las leyes de la naturaleza de las regularidades accidentales véase Ernest Nagel, *The Structure of Science, op. cit.,* capítulo 4, y el examen crítico de la posición de Nagel realizado por A. J. Ayer en *Scientific American* (junio 1961), 197-203. Véase también H. Reichenbach, *Nomological Statements and Admissible Operations* (Amsterdam: North Holland, 1954, y New York: Humanities Press, 1954). Un criterio completamente distinto acerca de las leyes de la naturaleza (la consideración de que son "principios de necesidad natural") puede verse en W. Kneale, *Probability and Induction* (Oxford: Oxford University Press, 1949), secciones 16-19.

4. LA EXPLICACION EN LAS CIENCIAS BIOLOGICAS

Las ciencias históricas

Gran parte de la biología es de carácter histórico: es *historia natural* en el sentido de que se ocupa de hechos particulares relativos al mundo terrestre que nos circunda. Parte de la biología implica incluso una *narrativa histórica*: yo estoy aquí considerando la historia de la evolución de las especies. Fuera de la biología también nos encontramos con ciencias que son esencialmente históricas. Por ejemplo, la geología se ocupa del estudio detallado de la corteza terrestre, y por tanto de un conjunto de hechos particulares. Por supuesto que la geología utiliza leyes generales de la física y la química para su explicación de estos hechos particulares, pero ello no le priva de su "status" de ciencia histórica, ya que su objetivo principal lo constituyen los propios hechos particulares y no los hechos particulares *en cuanto* comprobación de leyes físicas y químicas. También gran parte de la astronomía debe considerarse como historia: considérese, por ejemplo, la descripción detallada del sistema solar. Parte de la astronomía, sin embargo, debe ser considerada como física, es decir, teniendo como objetivo leyes físicas más que hechos astronómicos particulares. Como cuando las estrellas y otros cuerpos celestes se consideran como una especie de laboratorio natural que produce fenómenos que pueden ser utilizados como *comprobación* de teorías físicas. Cuando el objetivo primordial de la astronomía es el de la *cosmología* nos encontramos en una situación intermedia difícil de calificar. Considérese la explicación de la generación de energía en las estrellas: tal explicación se realiza mediante la aplicación de la física nuclear. ¿Se trata en este caso de "historia", es decir, de la explicación de hechos relativos a estrellas individuales, o de una parte de la historia natural relativa a determinados tipos de estrellas observados, o de la explicación de una verdadera ley de la naturaleza relativa a las estrellas de un tipo determinado? O, con otro ejemplo, ¿ha de considerarse la separación de unas galaxias de otras como parte de la historia natural relativa al universo, o se ha de considerar como una ley de la naturaleza relativa a las galaxias? En cosmología, al parecer, no están muy claros los límites entre historia natural y teoría científica. Confío, sin embargo, en ser capaz de establecer tal delimitación cuando nos ocupemos de la biología.

Historia natural y taxonomía en biología

Como introducción al estudio de la botánica o de la zoología se le dan algunas nociones al estudiante relativas a parte de la historia natural, es decir, se le ofrecen algunas generalizaciones acerca de los organismos que se encuentran en la superficie de la Tierra. Se le habla, por ejemplo, de las diferencias características entre los animales, las plantas, las bacterias y los virus, cómo se reproducen (ya bien sexual o asexualmente), si son células simples o si constan de multitud de células, qué se observa cuando las células se dividen y demás. Las nociones que el estudiante aprende de este modo son generalizaciones relativas a organismos terrestres, las cuales no se pretende que sean aplicables —al menos a este nivel— en todo tiempo y lugar. Por ejemplo, al aprender que las células de los animales y de las plantas contienen cromosomas, el estudiante no tiene que tomar en consideración la cuestión especulativa concerniente a si todos los animales y plantas que puedan existir en los planetas de las estrellas lejanas constan de células o si las células de tales animales y plantas lejanos contienen estructuras que se identifiquen con los cromosomas. El estudiante está aprendiendo únicamente parte de la historia natural terrestre. Es más complicada que la historia natural concerniente a los porteros de fútbol o a los "boy-scouts" (puede implicar indirectamente cierto conocimiento de la física y la química, como por ejemplo en el diseño de microscopios o en la utilización de tintes) pero sigue tratándose de historia natural al ocuparse de generalizaciones aplicables terrestremente más que de leyes universales de la naturaleza. Además, puesto que las plantas, los animales y las células vivas son cosas complejas y diversas, no hay razón alguna para que algunas de estas generalizaciones no tengan excepciones terrestres. Compárese la generalización de que los automóviles poseen radiadores [1]. Tal generalización resulta muy útil, aun cuando lleguemos a enterarnos de que hay excepciones a ella, como en el caso de que los coches se enfríen mediante aire. Tales excepciones carecen de importancia: al aprender cómo funciona un radiador no tenemos por qué preocuparnos del hecho de que en los coches sin radiador esta función se realice de modo diferente. Una excepción a una generalización de la historia natural no tiene trascendencia científicamente: nos obliga simplemente a cambiar "todos" por "casi todos". Así la generalización de que entre los animales que amamantan a sus crías ninguno pone huevos tiene que cambiarse simplemente por la generalización de que ningún animal que amamante a sus crías, excepto el "platypus" y el "echidna", pone huevos. Esto puede llevar, por supuesto, a una interesante investigación acerca de la economía interna del "platypus" y el "echidna", exactamente igual que alguien que tenga la creencia de que todos los coches poseen

[1] Utilizo aquí la palabra "radiador" con el significado de "radiador que hace uso de un líquido circulante".

radiador deseará echar un vistazo al interior de un automóvil que se enfríe mediante aire. De todos modos, las excepciones a las generalizaciones de la historia natural no nos cogen nunca de sorpresa, mientras que una excepción a una ley de la naturaleza putativa sí que resulta sorprendente y provoca serias dificultades.

Considérese la generalización de que las vacas tienen cuatro patas. Esto no significa que todas las vacas tengan cuatro patas. No hay motivo para que no nazca una ternera, que sea un fenómeno de la naturaleza, con tres o cinco patas. Por tanto, la generalización de que las vacas tienen cuatro patas no excluye tajantemente la suposición de que algunas vacas no tengan cuatro patas. De modo semejante la generalización de que los coches tienen radiadores no se ve afectada por la existencia de coches que se refrigeren mediante aire. Como tampoco le resta un ápice de eficacia a la explicación física de cómo funciona un radiador en un coche refrigerado mediante radiador el hecho de que algunos coches no tengan radiadores. En el mismo sentido, las explicaciones bioquímicas y biofísicas del funcionamiento de los organismos no quedan en absoluto invalidadas por el hecho de que no existan leyes acerca de los organismos, si no únicamente generalizaciones sin excepciones[2].

Supóngase que existiese una pareja vacuna con tres patas que, sorprendentemente, no produjese sino terneras con tres patas, y que de este modo obtuviésemos crías con tres patas. En este caso el enunciado de que las vacas no tienen siempre cuatro patas tendría una importancia taxonómica que no podría conseguirse por la mera existencia de algunos fenómenos de la naturaleza que presentasen tres patas. Quizá podríamos considerar a las vacas de tres patas como una sub-especie en una nueva etapa de evolución. Por supuesto que si la nueva raza de vacas con tres patas no se mezclase con las de cuatro patas probablemente reconoceríamos que esta raza constituía una especie distinta.

La clasificación de los seres vivos en especies (y más detalladamente en géneros, familias, órdenes, clases[3], grupos y reinos) tuvo su origen con anterioridad a la Teoría de la Evolución. Sin embargo, incluso en el período pre-evolucionario las características de los organismos que resultaban importantes para los taxonomistas generalmente no eran las mismas que las que llaman la atención del no erudito en la materia: en realidad son ca-

---

[2] No debemos concluir, de las consideraciones precedentes, la afirmación infundada de que estas generalizaciones no son expresables dentro de la notación de la lógica de predicados, complementada con un conjunto adecuado de predicados biológicos y de otro tipo. Lo que sí muestran, sin embargo, es que la formalización de "las vacas poseen cuatro patas" es considerablemente más complicada que "(x) x es una vaca $\supset$ x posee cuatro patas". Tenemos que poder formalizar la idea de que una cosa es miembro de una especie la mayoría de cuyos miembros poseen cuatro patas y esto podría resultar muy complicado, ya que la noción de especie implica la de las interrelaciones familiares y la capacidad de producir entre sí descendencia. La noción de especie también implica referencia a un particular, a saber, el planeta Tierra. Véanse las págs. 51-3 precedentes y las páginas 77-8 posteriores.

[3] En el sentido taxonómico, no teórico, de conjunto, por supuesto.

racterísticas de las que el no erudito no suele apercibirse. Debemos añadir ahora que una clasificación taxonómica es adecuada si y sólo si refleja apropiadamente las relaciones familiares de la especie en el árbol de evolución de la familia[4]. Cuando vemos que la clasificación en especies, géneros, etc., debe hacerse de modo que se adecúe a las hipótesis evolucionistas, entonces ya no parece esencial el que poseamos verdades universales de naturaleza semejante a la legal acerca de los objetos que así se clasifican, y no tiene por qué preocuparnos la posibilidad de vacas con tres patas (volviendo a nuestro ejemplo para andar por casa, bastante ingenuo desde el punto de vista biológico). Las vacas se definen por sus relaciones hereditarias con otras vacas (en su mayoría de cuatro patas) y no por la posesión de características como la de la posesión de cuatro patas. Existen, por supuesto, algunas características comunes a todas las vacas. Por ejemplo, todas poseen corazón y todas poseen huesos, ya que sin corazón y sin huesos no serían posibles. Con todo, si se pudiese dar el caso de extrañas vacas sin corazón o sin huesos, entonces el biólogo las estudiaría simplemente, y no se encontraría en la posición del físico que ha encontrado excepciones a leyes de la naturaleza que se venían cumpliendo con anterioridad.

De hecho, cualquier generalización de la biología descriptiva forma parte de la historia natural de la Tierra, y en la mayoría de los casos no será verdadera a menos que vaya acompañada de alguna referencia explícita o implícita a nuestro planeta. De lo contrario resulta fácilmente invalidada por algún fenómeno de la historia natural extraterrestre. Desde luego, la situación sería distinta si nuestros conocimientos fueran mucho mayores de lo que lo son en la actualidad. Supongamos que tuviésemos un conocimiento bioquímico completo del genotipo de una especie. Entonces es posible que pudiéramos (al menos si fuésemos mucho más listos de lo que somos en la realidad) deducir que los organismos que presentarán determinado genotipo (determinándose con exactitud el genotipo mediante un número de fórmulas químicas complejas, una por cada gen), desarrollándose en determinado ambiente físico y químico llegarían a desarrollar determinadas características fenotípicas. En este caso la proposi-

---

[4] Desde luego, estas relaciones no son siempre conocidas, especialmente en el caso de organismos prehistóricos, pero es de esperar que la clasificación taxonómica pudiera adaptarse al árbol familiar si éste fuese conocido en detalle. La observación hecha en el texto en realidad precisa de puntualizaciones a la luz de taxonomistas estadísticos tales como Robert R. Sokal [1], que insisten en la imposibilidad de llevar a cabo clasificaciones detalladas que se adapten al modelo evolucionista debido a que se conocen pocos detalles de la historia de la evolución para que esto sea factible. Por consiguiente, evitan las consideraciones de tipo evolucionista. Por lo que yo puedo colegir, su trabajo es importante, principalmente para producir un sistema clasificatorio que sería de máxima utilidad para lograr la reintegración de la información en las ciencias biológicas. Esto no parece oponerse al alegato de que una taxonomía evolucionista posee mayor interés teórico, aun cuando sea menos factible el realizarla con detalles. Debemos añadir, sin embargo, que Sokal ha mantenido también que a veces las clasificaciones que logramos mediante métodos estadísticos resultan ser importantes desde el punto de vista de la evolución.

ción de que los organismos con características genotípicas G que se desarrollen en un ambiente A tendrían características fenotípicas F sería un teorema de la física y de la química. Poseería ahora la universalidad cósmica de la que, como hemos visto, carecen las generalizaciones de la historia natural. Podría ser afirmada con relación a todo lugar y tiempo del universo, y podría ser deducida, desde luego, como proposición hipotética con independencia de que existiesen o no organismos del tipo en cuestión, v no haría referencia, explícita o implícita, al planeta Tierra.

Hemos estado considerando, en el parágrafo precedente, algo en realidad fantástico, ya que resulta dudoso que pudiera lograrse un conocimiento tan detallado ni siquiera en un futuro muy remoto. ¿Existe alguna proposición en biología, tal como hoy en día existe, que sea claramente biológica (y no únicamente una proposición de la física o de la química utilizada en explicaciones biológicas) que sin embargo no sea una proposición de la historia natural y que posea aplicabilidad cósmica? [5]. Volveremos a considerar esta cuestión más avanzado el capítulo.

La explicación en biología

Cuando una ciencia biológica está en su estadio más avanzado, sus explicaciones resultan ser bioquímicas o incluso biofísicas. Considérese, como un simple ejemplo, la cuestión de cómo las células vivas obtienen energía de la combinación de glucosa ($C_6H_{12}O_6$) y oxígeno (O_2). Si la glucosa se quema en una llama la reacción es simplemente

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O$$

En la conversión de glucosa y oxígeno en dióxido carbónico y agua se emite energía bastante violentamente en forma de calor y luz. Una célula necesita conseguir la energía a temperatura baja y a ritmo moderado. A fin de conseguir que la reacción antes mencionada tenga lugar a bajas temperaturas y de modo gradual, es necesario descomponer la reacción en una cadena de reacciones intermedias, cada una de las cuales es catalizada por ciertas encimas (moléculas proteínicas que actúan como catalizadores para determinadas reacciones específicas). La descripción y explicación detalladas de esta cadena de reacciones es evidentemente parte de la química orgánica. La *biología* se introduce simplemente mediante afirmaciones tales como que las células en cuestión contienen trifosfato de adenosina (TFA) y difosfato de adenosina (DFA), dos de las más im-

[5] Estoy utilizando "aplicabilidad cósmica" en un sentido bastante vago, indicando que la proposición será de interés para los científicos en todo tiempo y lugar. En el sentido *lógico* de "aplicable" son los predicados o las sentencias abiertas los que son aplicables, no las proposiciones. De hecho, la sentencia abierta "x es un tigre terrestre $\supset$ x tiene rayas" es verdadera con relación a alguna cosa, en todo tiempo y lugar, ya que algo que esté fuera de la Tierra no es un tigre terrestre.

portantes encimas implicadas en el proceso. Esto evidentemente forma parte de la historia natural, ya que no es preciso que nos preocupe a este nivel si las células de los planetas lejanos contienen TFA y DFA o cómo se metabolizan exactamente. (Por supuesto que este trozo de historia natural puede ser nuevamente explicado tal vez mediante otra parte de la historia natural —la que alude a que las células contienen ácidos nucleicos de tipos variados— y también más química.)

Es de gran utilidad el comparar una explicación biológica, tal como la que se discutió en el parágrafo precedente, con la explicación dada por un ingeniero respecto al comportamiento de una radio. En primer lugar se encuentra con el diagrama de cables, que establece que el aparato en cuestión tiene varios componentes: batería, espirales, condensadores, resistencias, válvulas, transformadores, altavoz, etc. Gran parte de la explicación se da en términos vulgares: decir que un cable une este componente con aquel otro es muy parecido a decir que una célula contiene determinadas divisiones que son visibles al microscopio. Parte de las descripciones del ingeniero son de un tipo más refinado, haciendo uso de algunos conceptos teóricos: por ejemplo, una válvula termiónica es algo con un filamento que emite electrones y un ánodo es algo que atrae a estos electrones —descripciones que sólo pueden ser comprendidas por una persona que posea algunos conocimientos teóricos—. Sin embargo, el que esta pieza en particular, que veo ante mis ojos, sea una válvula, es un enunciado particular. Podemos considerarlo comparable a un trozo de historia natural expresado científicamente, tal como "las células contienen TFA", o "los cromosomas contienen DNA" en contraste con las expresiones vulgares "los tigres tienen rayas" o incluso "esta célula contiene cromosomas".

Quisiera sugerir, por tanto, que al igual que las explicaciones relativas a la electrónica implican el uso de leyes físicas junto con proposiciones acerca de piezas de aparatos (establecidas, frecuentemente en forma de diagramas de cables), gran parte de la biología consiste en el uso de proposiciones de la física y la química junto con proposiciones de la historia natural (tanto de tipo refinado como vulgar). Por supuesto, existe una diferencia entre la electrónica y la biología en el sentido de que la primera es una ciencia aplicada, mientras que la biología, aunque tiene sus aplicaciones prácticas, se cultiva más bien para satisfacer la curiosidad intelectual. El ingeniero electrónico se ocupa principalmente de diseñar aparatos, y sólo en segundo término de explicarlos. Sin embargo, esto no invalida en lo más mínimo nuestra comparación de las explicaciones dadas en biología y en electrónica.

Podemos también fijarnos en un tipo de explicación que no es todavía bioquímica o fisiología, pero que lleva camino de ser bioquímica o biofísica. Por ejemplo, se puede explicar una enfermedad como motivada por la falta de alguna vitamina. Es evidente que este tipo de explicación presentaba algún valor incluso antes de que los conocimientos acerca de la química de las vitaminas fueran muy extensos. Se describía a las vitaminas

como sustancias que se encontraban específicamente en algunos alimentos y que eran necesarias para el funcionamiento del organismo de un modo que no se comprendía todavía. Las explicaciones pre-bioquímicas relativas a vitaminas, hormonas y demás, podrían ser comparadas con el diagrama esquemático de un ingeniero que no entra en detalles, pero que podrían ser proporcionados con más conocimientos. Por ejemplo, el diagrama esquemático (o diagrama "en bruto", como se le llama a veces) de un receptor de radio superheterodínico puede mostrar que la señal de la antena se alimenta de un amplificador de frecuencia, de ahí pasa a un modificador de frecuencia y de ahí a dos amplificadores de frecuencia intermedia en la serie, de ahí a un detector, de donde pasa a un amplificador de frecuencia auditiva y finalmente al altavoz. No especificará en detalles los circuitos de cada fase, pero las fases quedarán especificadas mediante la referencia a sus funciones dentro del aparato completo. La diferencia entre "amplificador de frecuencia intermedia" y el circuito en detalle de esta parte del receptor se asemeja bastante a la diferencia entre la expresión "vitamina C" y la fórmula química detallada del ácido ascórbico[6]. De modo semejante, alguien que no poseyese conocimientos acerca de la física del estado sólido podría considerar a los transistores simplemente como objetos que realizan las funciones llevadas a cabo previamente por las válvulas termiónicas: sólo tiene una idea muy ligera del *funcionamiento* del transistor, al igual que alguien que no conociese la bioquímica de la vitamina C tendría una idea muy ligera de su funcionamiento en el organismo, y lo único que sabría es que se trata de algo que el cuerpo humano necesita en cantidades muy pequeñas, que no es sintetizado por el organismo sino que ha de obtenerse del exterior, y sin el cual el cuerpo humano enferma en varios sentidos.

Las explicaciones teleológicas y evolucionistas en la biología

Como se observará, en la sección precedente se han utilizado algunas expresiones teleológicas. Por ejemplo: "necesario al organismo" o "necesario para evitar enfermedades". De modo semejante se utilizaron términos teleológicos dentro del contexto electrónico: decir que una parte de un aparato de radio es un "modificador de frecuencia" es especificar la parte mediante la función que realiza. Incluso un componente individual puede

---

[6] No del *todo*. El que el ácido ascórbico impida el escorbuto no puede verse en la fórmula química del ácido ascórbico, únicamente. Se necesitan conocimientos de la bioquímica del escorbuto. Por otra parte, el que una parte del circuito sea un amplificador de frecuencia intermedia puede verse observando esta parte del circuito únicamente. Una analogía más acertada podría ser la de un circuito en un receptor de televisión sintonizado para recibir determinada estación transmisora. La descripción funcional es "circuito sintonizado para recibir tal y tal estación". La explicación electrónica completa precisa *tanto* de las características eléctricas del circuito *como* de las del transmisor.

ser así especificado: para alguien que no conozca cómo se puede explicar el funcionamiento de un transistor, la palabra "transistor" no significará probablemente más que "cristal que contiene adulteraciones y electrodos unidos de tal y tal modo". Por supuesto que es posible que una persona no conozca ni siquiera todo esto en relación con la historia natural de los transistores, y para ella un transistor será definido funcionalmente como algo que rectifica o amplía señales.

Supóngase que preguntamos por qué un determinado receptor de radio posee un modificador de frecuencia (algunos no lo poseen). Se podría decir que el diseñador del receptor deseaba hacerlo sensible a las señales débiles. La amplificación en las frecuencias de radio e intermedias, con anterioridad a la fase del detector (en la que la señal se convierte en frecuencia auditiva) proporciona mayor sensibilidad que la amplificación en la fase de la frecuencia auditiva. La amplificación de la frecuencia de radio es difícil, ya que se dan fácilmente las inestabilidades. Por tanto, es habitual que no haya más que una fase de amplificación en la frecuencia de radio y que la mayor parte de la amplificación se realice en las fases de frecuencia intermedia. La respuesta, pues, a la pregunta de por qué los receptores de radio contienen frecuentemente modificadores de frecuencia radica en hacer observar que este método permite que se combine la amplificación con la sensibilidad. Presupone, por supuesto, que el diseñador del aparato *desee* que se combine la amplificación con la sensibilidad. Supóngase que contemos con una generalización perteneciente a la historia natural relativa a que la mayoría de los diseñadores de aparatos de radio desean a la vez gran amplificación y buena sensibilidad. Sería entonces fácil construir una explicación estadística o una explicación parcial según el modelo de Hempel (cf. el capítulo 3) acerca del hecho de que un determinado aparato contenga un modificador de frecuencia. Contendría proposiciones de electrónica (explicables ellas mismas mediante diagramas de cables y leyes de la naturaleza) y también una proposición acerca de lo que la mayoría de los diseñadores de aparatos de radio desean de éstos. (La explicación resultará ser sólo parcial si existen otros modos, aparte del modificador de frecuencia o superheterodino, para lograr una considerable sensibilidad junto con gran amplificación.)

Compárese esto con una explicación teleológica de biología. Supóngase que preguntamos por qué los animales de ciertos tipos tienen riñones. Puede alegarse como respuesta a esta pregunta que la finalidad de los riñones es remover los residuos químicos y regular la cantidad de agua de los tejidos. Carentes de riñones estos animales no podrían vivir. Hasta la época de Carlos Darwin la explicación putativa de la existencia de los riñones se remontaría a los propósitos de Dios: se presupondría que Dios deseaba que estos animales fuesen seres vivos y vigorosos además, al igual que el diseñador del aparato de radio deseaba combinar una sensibilidad considerable con una considerable amplificación. Tal referencia a los designios de Dios no proporcionaría una explicación científica en tanto en cuanto las proposiciones teológicas requeridas no formarían

(como es de suponer) parte de un sistema verificable mediante la observación. A partir de Darwin se ha hecho innecesaria la referencia a los designios de Dios, ya que en gran medida la teoría de la evolución explica por qué los organismos son, en gran parte, tal *como si* hubiesen sido concebidos según una finalidad. La evolución también explica, aunque parezca paradójico, por qué algunos organismos parecen haber sido *mal* concebidos en algunos aspectos. Por ejemplo, las cavidades del cuerpo humano no desaguan bien (como algunos hemos descubierto a nuestra costa), puesto que hemos evolucionado a partir de animales que caminaban a cuatro patas y con la cabeza hacia abajo en lugar de hacia arriba. La designación de las cavidades que resultaba una ventaja para la supervivencia de estos animales no resulta ventajoso para nosotros, pero sin embargo el proceso de evolución mediante mutación y selección natural no ha sido capaz de corregir esto y probablemente nunca lo podrá hacer. (Algunos cambios debidos a la evolución son irreversibles, por razones fáciles de comprender.)

La teoría de la evolución nos proporciona la explicación, o la explicación parcial, de casos de *aparente* designio, es decir, casos que desde un punto de vista pre-evolucionista (y quizás dentro de un tipo de teoría evolucionista como la de Lamarck) se considerarían debidos a un auténtico designio y a una auténtica finalidad, de forma muy semejante al modo en que las características de un aparato de radio se explican, parcialmente al menos, con referencia a los designios de su creador. También proporciona la teoría evolucionista las explicaciones o explicaciones parciales de casos que *parecen* deberse a un designio *equivocado* (como en el caso de las cavidades mencionado en el último parágrafo). A veces la explicación, o explicación parcial, de los designios *equivocados* aparentemente puede resultar muy sutil. Por ejemplo, se puede postular una conexión genética debida a que dos genes se encuentren situados en el mismo cromosoma. (La conexión es mayor si se encuentran próximos en el cromosoma, por lo que es más difícil que se separen mediante el *cruce* de cromosomas.) Una característica fenotípica desventajosa puede ser seleccionada a causa de su conexión genética con una característica *ventajosa* desde el punto de vista de la selección.

Debe observarse que aunque hablemos de la "teoría" de la evolución lo que se encuentra en los libros acerca de este tema es en su mayor parte *historia,* en el sentido de que se ocupa de seguir los pasos de la evolución de las especies terrestres: proceso histórico particular, aunque se trate de uno de gran duración e importancia. La paleontología es una ciencia evidentemente histórica. Normalmente no nos encontramos con explicaciones acerca de hechos paleontológicos, sino solamente con explicaciones parciales o esquemas de explicación. Por ejemplo, si la paleontología proporciona testimonios fósiles con relación a ciertas características fenotípicas puede conjeturarse que estas características tuvieron lugar a causa de que conferían alguna ventaja desde el punto de vista de la selección o porque en un estadio anterior de la evolución lo hacían, o porque estaban

ligadas genéticamente a algún tipo que es o fue en algún tiempo ventajoso para la selección. Esto puede considerarse del modo más natural como un esquema de explicación, una sugerencia para considerar posibles modos de explicación de la existencia de la característica en cuestión. Un esquema de explicación puede incluso ser formulado en la forma de una explicación estadística de Hempel:

La mayoría de las características fenotípicas se originan como consecuencia de una selección natural.
Esta característica es fenotípica.
... ...
Por consiguiente, probablemente esta característica se originó como consecuencia de una selección natural.

Sin embargo, una "explicación" tal no presenta mucho valor, y sería preferible considerarla simplemente como un esquema de explicación. Resulta muy vaga. Por ejemplo, la noción de "originarse como consecuencia de la selección natural" es muy amplia, si se deja que siga el camino circunvalatorio en el que a menudo opera la selección natural (cf. las observaciones hechas anteriormente acerca de las conexiones).

El carácter histórico de la teoría de la evolución se revela muy claramente en el hecho de que la aparición de los fenómenos que han de explicarse se descubre sólo a partir de *documentos históricos.* Por este motivo, aunque podemos ofrecer explicaciones, explicaciones parciales o quizá sólo esquemas de explicación del curso de la evolución en el pasado, no podemos predecir, normalmente, el futuro de la evolución. (*A veces* podemos predecir el futuro de la evolución. Cuando se inventa una nueva droga para combatir un tipo de bacterias es fácil aventurar que un tipo de estas bacterias que sea resistente a la droga evolucionará muy pronto.)

Considérese un sencillo caso de explicación evolucionista, el concerniente al melanismo de las polillas. Esto ha tenido lugar en las zonas industriales de Inglaterra desde hace siglo y medio. El humo de las chimeneas ha ennegrecido el follaje de los árboles, de modo que las polillas de color claro resultan más fácilmente visibles y por tanto son comidas por los pájaros. Ahora bien, con toda seguridad existe un gen para la tonalidad oscura en el grupo, puesto que siempre aparecen algunas polillas más oscuras entre las polillas de diversas especies, y en la mayoría de las zonas no industriales se ven desde luego antes y son comidas antes que sus parientes de color más claro. Por tanto, el gen de la tonalidad oscura queda normalmente *eliminado.* Sin embargo, en las zonas industriales son las polillas oscuras las que se encuentran en situación más ventajosa y es posible que sea el gen de la blancura el que quede eliminado del grupo.

Es evidente que si las conjeturas de las últimas sentencias fuesen formuladas con detalle la totalidad de la explicación podría ser realizada en una forma deductiva hempeliana. Lo que es interesante, sin embargo,

es que la única comprobación de algunas conjeturas evolucionistas puede ser la de que el hecho que ha de explicarse haya ocurrido. Quizá sea ilustrativa la comparación con el caso de caída de un puente que se explica mediante la hipótesis de fatiga del metal lo suficientemente intensa como para provocar la caída [2]. Es posible una explicación deductiva según el modelo hempeliano [3], pero esto se debe solamente a que la premisa general "todos los metales con el mismo grado de fatiga que el metal de este puente caen cuando sufren tal y tal presión" se alcanza con el testimonio independiente (de la memoria, las informaciones de la prensa, etc.) de que el puente en realidad cayó. De modo semejante las explicaciones deductivas de tipo evolucionista pueden ser formuladas según el modelo hempeliano, pues están basadas en las generalizaciones (que en mi opinión no son leyes de la naturaleza) de que ciertos genes son portadores de una determinada ventaja selectiva en una localidad dada en un tiempo dado, junto con otras generalizaciones, tales que hagan compatible la proporción en que los genes sufren mutación con la selección que de hecho tenga lugar [7].

Puesto que las hipótesis que se precisan en las explicaciones evolucionistas son a menudo conjeturas para las que el hecho que ha de ser explicado proporciona la principal prueba, nuestra capacidad para explicar el pasado excede con mucho a nuestra habilidad (casi nula) para predecir el futuro de la evolución. Tenemos *testimonios* del pasado (en su mayor parte testimonios fósiles) pero, no contamos con testimonios del futuro. (Esta última asimetría tiene una profunda explicación física que será discutida en el capítulo 8.)

El carácter histórico de las explicaciones evolucionistas o explicaciones parciales puede verse con suma claridad si consideramos la evolución del caballo de criatura con cuatro dedos a animal con un solo dedo. Esto tuvo lugar en dos momentos: de cuatro dedos a tres dedos, y mucho más tarde de tres dedos a un dedo [5]. Cada uno de estos momentos fue seguido de períodos más lentos de ajuste mecánico en los cuerpos de los animales, de tal modo que se adaptasen mejor al nuevo tipo de pie. Supongamos ahora que por una especie de milagro un biólogo que poseyese todo nuestro conocimiento moderno relativo a genética, bioquímica, etc., y también con una idea general de la evolución mediante la selección natural, viviese en la era prehistórica en la que el pie del caballo estaba a punto de evolucionar de cuatro dedos a tres. Es del todo evidente que sería totalmente incapaz de predecir el cambio de cuatro a tres dedos, y mucho menos todavía el modo en que (o incluso el hecho de que) el estado de tres dedos sería seguido del estado de un dedo. Podemos ofrecer algún tipo de explicación simplemente porque a partir de los documentos fósiles sabemos lo que ocurrió y por tanto podemos hacer conjeturas

[7] Han de considerarse a la vez el índice de mutación a genes que proporcionan la ventaja selectiva y el índice de mutación *hacia atrás*, hacia lo original, para lo cual quizá sea necesario el tipo de análisis estadístico que fue introducido por R. A. Fisher [4].

con buen fundamento acerca de los hechos biológicos necesarios para proporcionar la explicación, por ejemplo, de que hubo ciertas mutaciones de genes. Por si fuera poco, a partir de los testimonios geológicos, estamos en posesión de datos no biológicos, es decir, datos acerca de los cambios climáticos, que son necesarios para la explicación pero que no estarían a disposición de nuestro imaginario biólogo prehistórico.

Estas consideraciones no pretenden mostrar, en modo alguno, que las explicaciones evolucionistas no pueden ajustarse al modelo deductivo o estadístico hempeliano, como Scriven parece haber considerado. Sin embargo, cuando algunas pretendidas explicaciones evolucionistas se hacen ajustar al modelo hempeliano a menudo aparecen como explicaciones meramente parciales. Por ejemplo, a partir de determinadas conjeturas quizá podamos deducir que una especie adquirirá una de las características fenotípicas de una clase dada, pero no que presentará aquella particular que posee. De ser así, ello constituirá simplemente una explicación parcial de la posesión de la característica en cuestión.

Las explicaciones en la teoría de la evolución utilizan generalizaciones, pero éstas son generalizaciones de historia natural más que leyes de la naturaleza. Por ejemplo, el que todas las polillas (o una cierta proporción de ellas) de una determinada localidad y en un determinado período de la historia porten un gen recesivo de melanismo ha de clasificarse con toda certeza como un trozo de historia natural y no como una ley de la naturaleza. (Por las razones ofrecidas en las págs. 76-7.) De modo semejante, el historiador político o social utiliza generalizaciones (a menudo bastante falsas) acerca de ciertos tipos de seres humanos. Desde un punto de vista lógico la denominada "teoría" de la evolución no es teoría, en el sentido de un conjunto bien articulado de leyes de la naturaleza, sino historia. Darwin no se ocupó del mismo tipo de actividad que Newton, y no debe ser juzgado mediante los mismos criterios. Michael Scriven ha sugerido que el reconocimiento del carácter *histórico* de la teoría de la evolución podría ser de valor para los científicos sociales: las ciencias de que se ocupan no precisan de un segundo Newton, ya que el paradigma de Darwin muestra cómo puede proporcionar explicaciones un científico aun cuando el carácter de la materia de que se ocupa le impida ser capaz de ofrecer predicciones en la forma en que Newton podría hacerlas [6].

Las matemáticas en la ecología y en la teoría de la evolución

El carácter histórico de la teoría de la evolución no es incompatible con el hecho de que la teoría de la evolución pueda hacer uso de un sutil análisis matemático, como ocurre en la obra de Sir Ronald Fisher. Las conjeturas acerca de las frecuencias de los genes en los grupos, de sus frecuencias de mutación y demás son, de acuerdo con nuestros criterios, conjeturas de la historia natural, más que leyes de la naturaleza, pero no hay nada extraño en el hecho de que para descubrir sus implicaciones se

precise de análisis matemático complejo. La misma observación es válida con respecto a la ecología, que es el estudio de la relación de plantas y comunidades animales con su medio ambiente, y en particular las variaciones que tienen lugar en sus poblaciones. Supóngase que el índice de nacimientos de un cierto tipo de animales es uno determinado, el número de ellos que pueda vivir en cierta zona del país otro determinado y supongamos que contamos con estadísticas similares para otros tipos de animales (incluyendo los que vivan de sus presas) que tengan que vivir en el mismo país, entonces el análisis matemático nos permitirá calcular los cambios de población en los diversos tipos de animales. A veces la situación será tal que una clase de animales llegue a extinguirse y otras veces habrá oscilaciones: el aumento de presas dará lugar a más animales rapaces, lo que dará lugar a su vez a disminución de presas, lo que implicará disminución de animales rapaces, lo cual volverá a dar lugar a un aumento de presas, y así sucesivamente. También otras veces el análisis puede llevar a predicciones incluso más complicadas. Lo esencial, sin embargo, desde nuestro punto de vista, es que a partir de ciertas proposiciones de la historia natural (por ejemplo relativas al índice de nacimientos de una especie de canguros en una determinada localidad y en un cierto momento), se deducen otras proposiciones de historia natural (por ejemplo que no existirán tales canguros en tal lugar dentro de diez años). La matemática que se utiliza en toda ella es matemática pura, y las proposiciones no matemáticas que se utilizan son, por lo general, enunciados de historia natural y no leyes de la naturaleza

Biología cósmica

Aun cuando los detalles del proceso de la evolución en la Tierra sean una cuestión a estudiar como fragmento de la historia terrestre, podemos preguntarnos si existen, no obstante, algunas proposiciones generales acerca de la evolución que sean de carácter científicamente teórico, es decir, comparables a las leyes teóricas de la física en ser universales y aplicables a todo lugar y tiempo dentro del universo [8].

Parece ser bastante seguro que dondequiera en el universo que exista vida, los organismos vivos tendrán que poseer un código genético, compuesto de unidades discontinuas, al igual que el alfabeto contiene un número de unidades discontinuas (las letras), y consista en lo que consista este código tendrá que ser tal que la información codificada pueda ser reproducida exactamente o casi exactamente. Si se reproduce con exactitud quizá no quede del todo estable en el período que va del momento en que es reproducido al momento en que a su vez se reproduce a sí mismo. Es como si hubiese de copiarse un documento con gran exactitud, una y otra vez, sólo con algunos errores aislados de cuando en

[8] Tomando el término "aplicable" con las debidas precauciones, cf. la nota 5.

cuando. En los organismos terrestres el código consta de moléculas de ácido deoxiribonucleico (DNA). Cada molécula de DNA consta de una cadena de amino-ácidos, de los que existen 20 tipos, y el "alfabeto" del código consta de tríos de estos amino-ácidos. Los errores que se dan o bien durante la reproducción o entre reproducciones del código se identifican con las mutaciones que fueron postuladas en la genética. No debe considerarse extraño el que tengan lugar tales mutaciones: quizá lo que debiera llamar más la atención es la gran *estabilidad* de las moléculas de DNA, y el que los cambios que se dan en ellas tengan lugar tan escasamente como tiene lugar. (¡Un libro sin erratas de imprenta es mucho más extraño que un libro con erratas de imprenta!)

¿Es cierto con relación a las criaturas vivas de todo tiempo y lugar del universo que su código genético tiene que constar de DNA? Lo sea o no, parece plausible el que en todo tiempo y lugar las criaturas vivas deben poseer un código genético u otro, y que este código no será del todo estable. En cuyo caso parece plausible que la siguiente proposición sea verdadera en todo tiempo y lugar: "Siempre que en el universo exista vida y exista un medio ambiente cambiante tendrá lugar la evolución" [9]. Esto, en realidad, parece asemejarse a una ley biológica de la naturaleza aunque sea un poco vaga en su formulación. Se basa en el hecho de que según lo que conocemos acerca del modo en que las criaturas vivas pudieron probablemente haberse originado a partir de la materia inerte, deben funcionar de acuerdo con principios físicos (o químicos) y deben incorporar un código genético de auto-reproducción. Así pues, quizá la proposición acabada de mencionar sea una de las pocas afirmaciones biológicas que pueden ser consideradas como auténticas leyes de la naturaleza, de alcance absolutamente universal, y que no contienen ningún elemento de historia natural. Con todo, no es *importante* para el biólogo en el sentido en que una ley física puede ser importante para un físico: un biólogo podría negarse a aceptarla y sin embargo proseguir perfectamente su trabajo ordinario, que consiste en explicar el comportamiento de los organismos *terrestres*. Después de todo, la proposición podría ser puesta en duda por un biólogo que dudase de que en todas partes del universo la vida poseyese la característica de la mortalidad.

¿Existen algunas otras posibles leyes que apliquen en todo tiempo y lugar a los objetos vivos? Un posible candidato quizá fuese: "Toda la materia viva contiene DNA." Pero, ¿qué significa "materia viva"? ¿Qué decir de las máquinas auto-reproductoras [7] que pudieran haber sido construidas por seres vivos? Quizá podrían incluso esclavizar a los seres que

---

[9] Es necesario el requisito de medio ambiente cambiante, porque aunque a menudo se da evolución en un medio ambiente estático, no hace falta que ocurra así, ya que los organismos pudieran estar ya adaptados del mejor modo posible a este medio ambiente (es decir, cualquier mutación que pudiese darse conduciría a una peor adaptación). Aun cuando el medio ambiente físico sea estático, puede haber evolución, por supuesto, si existen dos especies de organismos que se interrelacionen, ya que cualquier cambio que se dé en uno constituye un cambio en el medio ambiente del otro.

las habían construido y evolucionar ellas mismas mediante selección natural (su código podría estar en una cinta, en lugar de en DNA, y las erratas de imprenta ocasionales constituirían las mutaciones). ¿Qué motivo habría para negar el epíteto de "vivos" a tales seres? Si se objeta que los seres vivos no son *construidos* nunca esto resultaría una arbitrariedad. Se solía pensar que Adán había sido construido por Dios y esto no parecía considerarse una razón suficiente para negar la cualidad de "vivo" a Adán o sus descendientes. Sin embargo, supongamos que pudiera darse una determinada definición de "vivo" que negase esta cualidad a las máquinas auto-reproductoras. Quizá un modo de hacerlo sería estipular que el código genético debe estar realizado con la estructura de *moléculas*. ¿Sería entonces verdadero que todos los seres vivos contienen DNA? Esta cuestión depende evidentemente en parte de otras cuestiones, tales como ¿están todos los sistemas vivos basados en compuestos carbónicos? Quizá existan sistemas basados en compuestos de silicona en vez de carbono, y evidentemente éstos no contendrían DNA, que es un compuesto de carbono. Suponiendo que existan fundamentos para pensar que los compuestos de silicona no podrían poseer la variedad requerida y la complejidad necesaria para formar sistemas vivos, queda entonces todavía por resolver la cuestión de si podrían existir compuestos de carbono distintos de DNA que presentasen las propiedades requeridas de auto-reproducción. ¿Tenemos conocimiento de esto, o cuando menos razones para suponerlo?

Echemos un breve vistazo a la cuestión acerca de si todos los sistemas vivos del universo están basados en carbono. Pueden darse razones para suponer una respuesta afirmativa a esta cuestión. Debido a sus cuatro valencias el carbono puede, especialmente en combinación con el hidrógeno, formar cadenas y anillos, y por tanto puede dar lugar a la formación de moléculas muy complejas. La silicona tiene cuatro valencias también, y por tanto se han hecho especulaciones acerca de posibles sistemas vivos basados en silicona. Con todo, aunque la silicona presente prometedoras propiedades químicas probablemente no tenga las propiedades físicas [10] adecuadas. A temperaturas ordinarias sus compuestos son sólidos duros: por ejemplo, la roca de cuarzo dura, que es dióxido de silicona. Las células vivas tienen que constar de líquido en gran parte, a fin de que puedan tener lugar las complejas reacciones químicas. Al parecer los sistemas vivos basados en silicona tendrían que existir en un medio ambiente en el que el cuarzo fuera líquido, o quizá incluso gaseoso. En nuestro planeta, por ejemplo, la posibilidad de vida depende del hecho de que el dióxido de carbono exista en forma gaseosa a una temperatura a la cual es soluble en el agua, estando el agua en forma líquida. ¿Existirían gases y líquidos análogos asociados con la silicona a la temperatura a la que el cuarzo es líquido? También ha sido sugerido que una cadena de nitrógeno-fósforo-cloro sumergida en sales disueltas podría formar la base de los sistemas vivos [8]. Mas, como W. T. Williams ha

[10] *Físicas*, en el sentido de opuestas a *químicas*.

señalado [9], el problema que estas especulaciones entrañan es que uno se encuentra con que tiene que hacer un montón de ellas al mismo tiempo, ensamblándolas todas juntas adecuadamente. Por este motivo, señala Williams, "los mundos químicamente extraños, aunque son concebibles, tienden a ser no plausibles". Por ejemplo, no valdría postular un sistema vivo dependiente de amoníaco líquido, que pudiese existir únicamente en un planeta frío como Júpiter, y también dependiente de óxidos de silicona líquida, que sólo existirían en un medio ambiente extremadamente cálido. Bien pudiera ser que el único sistema y medio ambiente en el que todas las cosas encajen entre sí correctamente sea un sistema de carbono en un medio ambiente químico o físico como el que se da en nuestro planeta Tierra.

Aunque fue publicado hace más de cincuenta años, el libro de L. J. Henderson, *The Fitness of the Environment* [10], que trata detalladamente de las combinaciones particulares de los accidentes afortunados producidos por nuestro medio ambiente terrestre, es una excelente fuente de consideraciones relevantes a las cuestiones que hemos estado discutiendo en el último parágrafo. Por supuesto, no es de mi competencia el discutir estas cuestiones y las observaciones acabadas de hacer tienen sólo por objeto sugerir los posibles modos en que los científicos podrían discutir en pro y en contra de la suposición de que el universo contiene sistemas vivos dependientes de una base química muy distinta a la nuestra. En cualquier caso esto servirá para indicar la dificultad de formular leyes de biología que pudieran ser aplicables en todo tiempo y lugar del universo.

Psicología

Vale la pena considerar hasta qué punto es aplicable la posición adaptada en este capítulo a la psicología. Después de todo, la psicología se ocupa de la explicación del comportamiento humano y animal, y por ello debe ser considerada con toda seguridad como una rama de la biología. La completa comprensión del comportamiento humano o animal, desde un punto de vista científico, parece depender de la comprensión completa de la estructura neurofisiológica del cerebro humano o animal. Es ésta una inmensa tarea, especialmente teniendo en cuenta que parece haber razones para suponer que el circuito detallado de dos cerebros cualesquiera es distinto en cada uno de los casos, al igual que las huellas dactilares de dos seres humanos cualesquiera no son del todo semejantes. Los seres humanos no se producen como modelos semejantes a los receptores de radio de una marca determinada, aunque incluso en el caso de los aparatos de radio nos encontramos con que algunos modelos que parecen ser exactamente iguales se comportan de un modo desconcertantemente distinto, debido a imperceptibles diferencias no fácilmente detectables.

Siguiendo una analogía con la electrónica, a primera vista pudiera parecer, por consiguiente, que las explicaciones detalladas de la conducta humana y animal no serían posibles sin un conocimiento exacto del cir-

cuito neurofisiológico (así como otras cosas, por supuesto, tales como conocimiento relativo a la producción de hormonas de diversas glándulas). Consecuentemente parecería quizá que al igual que distamos mucho de poseer conocimientos del circuito detallado de los seres humanos y de los animales a nivel individual también distamos mucho de haber alcanzado explicaciones psicológicas, al menos de un tipo científicamente respetable. Sin embargo, si consideramos una vez más la analogía con la electrónica podremos ver razones para nuevas esperanzas. Como se recordará una explicación electrónica no siempre entra en detalles, si no que se mantiene al nivel más abstracto del diagrama en bruto. De ese modo señalará, por ejemplo, que la salida de un amplificador de frecuencia de radio se conecta con la fase de modificador de frecuencia, y de ahí pasa a un amplificador de frecuencia intermedia, y así sucesivamente. Deja sin determinar en qué consiste el circuito detallado de estas fases, por ejemplo, si la amplificación tiene lugar mediante válvulas termiónicas o mediante transistores y si el modificador de frecuencia consta de dos válvulas separadas o de una válvula con muchos electrodos. Quizá, pues, será mejor considerar algunas de las "variables intermedias" (intervening) o "construcciones hipotéticas" postuladas por los psicólogos ya bien como elementos de una explicación tipo "diagrama en bruto" del funcionamiento del organismo (especialmente de su circuito neuronal) o como propiedades, estados o actividades de tales elementos. Es decir, los conceptos del psicólogo, tales como "impulso" (drive), "inhibición", "necesidad", "creencia", etc., son comparables a los conceptos electrónicos de "amplificador de frecuencia intermedia", "oscilación", "regresión", etc., y no a los de "transistor", "resistencia de 1.000 ohmios", etc. [11].

El diagrama en bruto de un receptor de radio proporciona la explicación de su funcionamiento, aunque no sea una explicación tan profunda como la del diagrama de cables completo. Es decir, explica cómo sale al exterior, a través del altavoz, el sonido correspondiente a la frecuencia que se imprime mediante la modulación de la onda transmisora con la que está conectada la fase de frecuencia de radio. Debe tenerse en cuenta que proporciona solamente una explicación *parcial* del hecho, por ejemplo, de que los sonidos *fuertes* [12] salgan por el altavoz. Para que proporcionara más datos tendríamos que saber no sólo la fuerza de la señal que entra y la colocación del control de volumen del sonido, sino también características detalladas del circuito, que no son proporcionadas por el diagrama en bruto. De modo semejante una explicación psicológica, aun cuando explique un modo de conducta (tal como una manifestación de ira) con toda probabilidad proporcionará únicamente una explicación *parcial* de un *movimiento* particular (tal como el puño de una persona moviéndose

[11] Considérese la semejanza con un diagrama en bruto de un ingeniero electrónico de los diagramas de la obra de J. A. Deutsch *The Structural Basis of Behavior* [11].

[12] Y todavía explica menos que los sonidos de "Waltzing Matilda" salgan al aire.

con rapidez y golpeando la nariz de alguien). El concepto de *conducta* es en cierto modo más abstracto que el de movimiento y en cierto modo menos abstracto. Es decir, muchos movimientos distintos pueden referirse a un tipo único de conducta. Usando el ejemplo proporcionado por J. A. Fodor [12], el correr hacia la derecha y el nadar hacia la derecha en un laberinto de forma *T* pueden considerarse como el mismo tipo de comportamiento (volverse hacia la derecha) dentro del contexto de una explicación psicológica. Es decir, la explicación puede referirse al movimiento de vuelta hacia la derecha realizado por un organismo y no ocuparse de si es cuestión de correr o nadar. Por otra parte, muestras del mismo movimiento pueden remitir a distintos tipos de conducta si lo que las motiva en el sistema nervioso central (o en cualquier otra parte) es diferente. Por ejemplo, si alguien le da en las narices a un conocido a causa de un movimiento reflejo esto no se refiere al mismo tipo de conducta (la agresiva, por ejemplo) que un golpe en la nariz originado por la ira y el deseo de agredir al conocido. La distinción entre movimiento y conducta no es por tanto en modo alguno incompatible con la consideración mecanicista de la psicología humana y la psicología animal.

Esto quedará aclarado, probablemente, si consideramos un artefacto aéreo sin piloto que sea concebido para aterrizar en algún punto distante mediante un mecanismo intencional (purposive)[13]. Es puesto en ruta, supongamos, pero el mecanismo del cohete asegura su vuelta a la ruta debida. Es decir, el artefacto contiene receptores y un computador que permiten que la intensidad y dirección de los desvíos del artefacto de su ruta correcta sean representados mediante cantidades físicas que se conectan con otro mecanismo del artefacto que hace volver al artefacto a su ruta de nuevo. Comparemos la vuelta del artefacto a su ruta causada por el mecanismo intencional, por una parte, y el que su vuelta tenga lugar por la acción de una ráfaga de viento descarriada, por otra parte. El primer caso corresponde a la conducta (golpear a una persona en la nariz), mientras que el segundo corresponde a un "simple movimiento" (dar con el puño de uno, accidentalmente, en la nariz de otra persona).

La explicación de la conducta es, en algunos aspectos, más difícil y, en otros, más fácil que la explicación de simples movimientos físicos. Así, los movimientos de tipo reflejo pueden ser explicados mediante generalizaciones de historia natural tales como "un golpe en la rodilla de la mayor parte de los seres humanos hace que el pie se levante ligeramente", a partir de la cual no es demasiado difícil una explicación en términos neurofisiológicos. Por otra parte, los movimientos particulares implicados en una muestra de conducta pueden resultar más difíciles de explicar que la propia conducta. Puede haber una explicación psicológica del tipo de diagrama en bruto del hecho de que me vuelva hacia la izquierda, hacia una calle lateral donde sé que puedo tomarme una cerveza, pero no de

[13] Una discusión de los mecanismos intencionales puede verse en el capítulo 9 posteriormente.

que me vuelva hacia la izquierda dando unos pasos de exactamente 30 pulgadas y media, y produciendo simultáneamente un sonido con mis dedos.

B. A. Farrell, en un interesante artículo "On the Limits of Experimental Psychology" [13] establece de modo persuasivo que no hay esperanzas de llegar muy lejos en psicología experimental simplemente mediante la búsqueda de leyes (o generalizaciones, como yo diría) del tipo estímulo-respuesta. Los estados internos y los procesos del sistema nervioso del animal son decididamente importantes, y por ello no puede darse ninguna sola generalización que relacione el estímulo con la respuesta. Cada nuevo estímulo cambia el estado interno del animal, y los propios estados internos pueden causar cambios en estados internos. Farrell sugiere que sigamos el ejemplo del físico y que tratemos de explicar los hechos relativos a lo mayor en términos de hechos acerca de lo más pequeño. En el caso de la explicación de la conducta esto significa descender al nivel neuronal o incluso más abajo. Y, por supuesto, existen psicólogos experimentales cuyo trabajo se realiza desde un punto de vista más bien fisiológico. Otros psicólogos que, al igual que Farrell, rechazan el método de estímulo-respuesta puro, o conductista, se sienten sin embargo más escépticos con relación al valor de las consideraciones fisiológicas. Uno de los principales exponentes de este punto de vista, J. A. Deutsch, mantiene que en el presente estado de desarrollo de la ciencia fisiológica en cualquier caso es mejor proceder mediante la postulación de elementos que presentan una determinada estructura abstracta o que realizan determinadas funciones definidas en términos abstractos. Consideremos, como él propone, la clarividente máquina de aprender descrita en su libro [14]. Cuando la máquina aprende algo cambia su estado en cierto modo, pero cómo esto ocurra exactamente carece de importancia, con tal de que se dé el cambio de un mecanismo apropiado de un estado adecuado a otro. En un compuesto electrónico de la máquina tal elemento podría ser un selector dekatron, y en otro electro-mecánico podría ser realizado por un relé que se auto-sostenga. También podría ser realizado mecánicamente por un brazo uniselector que se pusiese en reposo. Ahora bien, es evidente que podemos hablar acerca de tal mecanismo de cambio sin preocuparnos de su constitución física. Puede ser descrito de forma abstracta, o en términos de su función, aun cuando no conozcamos los detalles de sus realizaciones físicas. Igualmente puede ocurrir que contemos con buenos testimonios empíricos para una explicación estructural abstracta de la actuación de un animal o un hombre, tratándolo como a "una cámara oscura" cuyos estados estructurales internos conjeturamos a partir de su conducta manifiesta, y sin embargo podemos carecer de fundamentos sólidos para especular acerca de la actuación neurofisiológica de tal estructura abstracta. De igual modo muchas estructuras neurofisiológicas distintas pueden funcionar de modo parecido, y por tanto el método de la cámara oscura da lugar a una simplificación muy útil, en tanto en cuanto

los animales pueden ser tratados como si fuesen más parecidos internamente de lo que quizá lo sean.

Deutsch rechaza las especulaciones fisiológicas dentro de la psicología por razones metodológicas [15]. Señala que si a una hipótesis estructural para la que existen buenas pruebas se le da un planteamiento fisiológico que sea bastante especulativo, el abandono de la fisiología en un estadio posterior de la investigación puede llevar a los psicólogos a abandonar erróneamente la estructura abstracta de la hipótesis. También alega que una explicación psicológica cobra un aire de falsa precisión cuando se identifican los elementos del mecanismo explicativo con entidades fisiológicas, por lo que puede ser pasada por alto la falta de solidez de la parte estructural de la explicación.

Las consideraciones acabadas de mencionar son evidentemente de tipo táctico. Deutsch en modo alguno rechaza la ayuda de las consideraciones fisiológicas. Admite que a partir de las observaciones fisiológicas podemos obtener pruebas o al menos pistas relativas a la estructura de los mecanismos psicológicos. Como tampoco nada de lo que Deutsch dice excluye la posibilidad de explicaciones fisiológicas detalladas dentro de la psicología (si bien la propia complejidad del sistema nervioso haga que sean poco probables). La posición del que proporciona diagramas en bruto dentro de la electrónica no es en modo alguno incompatible con la esperanza de proporcionar circuitos detallados en algún otro nivel de la investigación. A veces, sin embargo, incluso en electrónica una explicación física no precisa de conocimientos relativos al circuito en detalle. Considérese, por ejemplo: "La radio se ha parado debido a que las pilas se han gastado." De modo semejante podría darse una explicación relativa a la falta de inteligencia de una persona mostrando una deficiencia en el tiroides.

Farrell, por su parte, a la vez que sugiere que podría ser buena estrategia el descender al nivel neuronal o incluso más abajo, también advoca explícitamente por investigaciones del tipo de las de Deutsch (a quien menciona), a saber, conjeturas y comprobaciones de elementos estructurales del sistema nervioso a gran escala. (Es importante recordar, al propósito, que una pieza de un aparato puede funcionar como dos elementos distintos de la explicación estructural: considérese una línea telefónica transportando conversaciones en diversas frecuencias de canal. En un diagrama de comunicaciones aparecerá como diversas líneas diferentes.)

En un valioso artículo en el que propugna una posición relativa a las explicaciones psicológicas que es bastante parecida a la de Deutsch, J. A. Fodor [16] señala con suficiente claridad que una función no es un trozo de material duro o blando. Esto es absolutamente cierto, mas los elementos de la explicación psicológica pueden ser elementos estructurales definidos abstractamente por su función o (si se desea, el evitar la semejanza con una teleología no deseable) por *lo que ellos hacen.* Fodor proporciona una analogía con la mecánica. Considérese la expresión funcional "elevador de válvula" y la expresión mecánica "mango" (cam-

shaft). En una explicación funcional puede dejarse sin decidir el que la función de elevar la válvula sea realizada mediante un mango o cualquier otro medio. "Un impulso no es un estado neurológico", dice Fodor, "tiene el mismo estatus lógico que la observación 'un elevador de válvula no es un mango'" [17]. En esta expresión, creo yo que el lenguaje de Fodor puede dar lugar a algunas malas interpretaciones. Algunos mangos (si se han conectado adecuadamente con las válvulas) son en realidad elevadores de válvulas, y de modo semejante es, con toda seguridad, legítimo conjeturar que algunos estados neurofisiológicos son en realidad las cosas que se describen como impulsos en la descripción más abstracta de los psicólogos. Las frases "un impulso no es un estado neurológico" y "un elevador de válvula no es un mango" pueden dar lugar a malos entendidos: pueden sugerir que todos o algunos de los impulsos son algo distinto a estados neurológicos y que todos o algunos de los elevadores de válvulas no son mangos, mientras que lo que Fodor realmente quiere decir es más bien que el *concepto* de impulso no es lo mismo que el *concepto* de estado neurológico y que el *concepto* de elevador de válvula no es lo mismo que el *concepto* de mango. Y esto es evidentemente cierto e importante.

Las explicaciones de la psicología, ya sean de tipo conductista, estructural o neurofisiológico, se basan todas ellas en lo que, desde el punto de vista adoptado en este libro, no constituye leyes en sentido estricto, sino generalizaciones de la historia natural. La generalización de que un estímulo determinado aplicado a un determinado tipo de animal en determinadas situaciones da lugar a un tipo de respuesta determinada se refiere evidentemente a ese tipo de animal terrestre (e incluso en este caso contará con excepciones en el caso de individuos anormales). De modo semejante, las hipótesis relativas a los elementos estructurales o el circuito fisiológico detallado constituyen generalizaciones de historia natural inferida, más que leyes. Con todo, las "cámaras oscuras" son construidas con frecuencia de tal modo que se comportan de modo bastante regular, y así, en algunos aspectos, los organismos pueden ser muy regulares en su comportamiento. La selección natural ha tenido esto en cuenta. Si consideramos cuántas veces hemos cruzado una calle llena de gente y vehículos, o hemos evitado que un balón en movimiento rápido nos diese en la cabeza, podremos comprobar la semejanza que guardan con las leyes algunos de los aspectos de nuestra conducta. Esto explica por qué algunos psicólogos se han inclinado a considerar a la psicología más parecida a la física de lo que en realidad es; mas sus relaciones con la biología y la analogía con la electrónica deben servir para corregir de algún modo tal impresión.

Sugerencias respecto a lecturas posteriores

Las diferencias entre las leyes de la naturaleza y las generalizaciones de la historia natural pueden verse en Stephen Toulmin, *The Philosophy of Science* (London: Hutchinson, 1953, y en rústica, New York: Harper and Row, 1960), especialmente págs. 44-56. Sobre el carácter histórico de la teoría de la evolución, el artículo de Michael Scriven "Explanation and Prediction in Evolutionary Theory", *Science,* CXXX (1959), 477-482, resulta muy sugerente y lleno de aciertos. Sobre la metodología dentro de la psicología véase J. A. Deutsch, *The Structural Basis of Behaviour* (London: Cambridge University Press, 1960, y Chicago: Chicago University Press, 1960), capítulo I; Jerry A. Fodor, "Explanation in Psychology" en Max Black (eds.), *Philosophy in America* (London: Allen and Unwin, 1965 e Ithaca, N. Y.: Cornell University Press, 1965), págs. 161-179; y B. A. Farrell, "On the Limits of Experimental Psychology", *British Journal of Psychology,* LVI (1955), 165-177. Compárese con las observaciones escépticas de la obra de L. Wittgenstein, *Philosophical Investigations* (New York: Macmillan, 1953), página 232 y también con el método propiamente conductista propugnado por B. F. Skinner. Véase, por ejemplo, la obra de Skinner, *Science and Human Behaviour* (New York: Macmillan, 1953, y en rústica, New York: Free Press, 1965). Una aguda crítica a la posición conductista es la proporcionada por N. Chomsky en su comentario a la obra de Skinner, *Verbal Behaviour* en *Language,* XXXV (1959), 26-58. Uno de los principales propulsores del punto de vista fisiológico es D. O. Hebb. Véase su *Organization of Behaviour* (New York: Wiley, 1961), y algunas modificaciones en su trabajo "Alice in Wonderland, or Psychology among the Biological Sciences" en H. F. Harlow y C. N. Woolsey (eds.), *Biological and Biochemical Bases of Behaviour* (Madison: University of Wisconsin Press, 1958).

Obras de tipo general sobre la filosofía de la biología incluyen la de J. H. Woodger, *Biological Principles* (London: Kengan Paul, 1929, y New York: Humanities Press, 1966) y la de Félix Mainx, "Foundations of Biology", *International Encyclopedia of Unified Science* (Chicago: University of Chicago Press, 1955). F. A. Hayek ha escrito un interesante artículo, "The Theory of Complex Phenomena" en Mario Bunge (ed.), *The Critical Approach to Science and Philosophy, In Honor of Karl Popper* (New York: Free Press, 1964). Sobre el tipo de explicación teleológica dentro de la biología véase Ernest Nagel, *The Structure of Science* (New York: Harcourt Brace and World, 1961), capítulo 12, y R. B. Braithwaite, *Scientific Explanation* (London: Cambridge University Press, 1953, y en rústica, New York: Harper and Row, 1960, traducción española: *La explicación científica,* ed. Tecnos, Madrid, 1964, capítulo 10. Acerca de cuestiones relativas a la vida en otros mundos véase el artículo de W. T. Williams, "Problems of Alien Biology", *Humanist,* LXXIX (noviembre 1964), 329-332; L. J. Henderson, *The Fitness of the Environment* (New York: Macmillan, 1913, y en rústica, Boston: Beacon Press, 1958); A. I. Oparin, *The Origin of Life,* traducido por Sergius Morgulis, segunda edición con una nueva introducción a cargo del traductor (New York: Dover, 1953) traducción española: *Origen de la vida sobre la Tierra* (ed. Tecnos, Madrid, 1970); y George Wald, "The Origin of Life" en *The Physics and Chemistry of Life* [A Scientific American Book] (New York: Simon and Schuster, 1955). La obra de Merle B. Turner, *Philosophy and the Science of Behavior* (New York: Appleton-Century-Crofts, 1965), es una revisión bastante completa de la filosofía y la metodología de la psicología experimental y teórica y contiene amplia bibliografía.

5. LEYES Y TEORIAS EN FISICA

LEYES EXPERIMENTALES, MEDIDA Y APROXIMACIÓN

Algunas leyes de la ciencia son aceptadas a la base de observaciones experimentales incluso antes de que éstas (o algo muy parecido a ellas) puedan ser deducidas de algún cuerpo de teoría bien articulado. Considérese la ley de Ohm. Esta puede ser explicada ahora a la base de la física del estado sólido, pero la derivación no es en modo alguno sencilla, y sólo pudo ser hecha en los últimos años. En su mayor parte, la ley de Ohm ha sido, y todavía es, aceptada como una generalización a partir de observaciones experimentales. Supóngase que se toman en cuenta la diferencia potencial a través de una resistencia E y la corriente a través de la resistencia I, y que los resultados se marcan en un gráfico. Se descubrirá que los puntos así marcados forman aproximadamente una línea recta. A causa de errores experimentales los puntos no formarán *exactamente* una línea recta, pero la línea recta será adecuada, de tal modo que ninguno de los puntos que estén más lejos que ella pueden ser explicados mediante errores experimentales. Por supuesto, podríamos obtener una curva exacta trazando una curva en zig-zag, u ondulatoria a través de todos los puntos marcados, pero esto no proporcionaría ninguna ventaja, por dos razones: Hacerlo supondría presuponer que existe siempre un error experimental cero, lo cual es improbable. Por supuesto, el conjunto exacto de errores experimentales implicados por la línea recta es en principio igualmente improbable, salvo que contamos con un prejuicio a favor de la simplicidad, y, por consiguiente, consideramos a la línea recta más probable que la línea en zig-zag u ondulatoria. Otra razón para considerar a la línea recta ilustrativa de la ley, más que la línea en zig-zag o la ondulatoria, es que si realizamos más observaciones nuestros nuevos puntos marcados no aparecerán, en general, en nuestra línea en zig-zag o nuestra línea ondulatoria original. Podríamos, igualmente, trazar la curva más suave posible consistente con el margen de error experimental conocido (o conjeturado).

Las leyes puramente experimentales se establecen a menudo en forma de polinomio, ya que cualquier función matemática continua cuyos primero, segundo, tercero, etc., derivados existan, puede ser representada aproximadamente mediante un polinomio con un número finito de términos. (Considérese, por ejemplo, la ley de la dependencia de la resistencia eléctrica en la temperatura.) Siempre se puede lograr mayor exactitud aumentando el número de términos. Así pues, supóngase que en un principio las observaciones experimentales sugieren la ley "$y=a_0+a_1x$". A la

luz de experimentos más exactos quizá resulte más idónea una ley tal como "$y=a_0+a_1x+a_2x^2$". Posteriormente experimentos aún más exactos pueden llevarnos a la ley "$y=a_0x+a_1x+a_2x^2+a_3x^3$", y así sucesivamente, aunque, generalmente, este proceso no continúa muy lejos.

A veces, sin embargo, será adecuada alguna función matemática bien conocida (dentro de los límites del error experimental) y, por tanto, puede ser propuesta una ley tal como "$y=\operatorname{sen} x$" o "$y=\log x$" o "$y=e^x$". Puede ser, incluso, que sean conjeturadas como leyes experimentales algunas funciones en apariencia muy complicadas: por ejemplo, la ley de Snell en la óptima geométrica es de la forma:

$$\text{“}y=\operatorname{sen}^{-1}(\text{constante}\times\operatorname{sen} x)$$

Sin embargo, la ley es, de hecho, más atractiva de lo que parece aquí, como puede verse considerando el modo habitual de escribirla:

$$\frac{\operatorname{sen} y}{\operatorname{sen} x}=\text{constante}$$

Es evidente, sin embargo, que en ausencia de una derivación teórica es decir, si la ley en un momento determinado del desarrollo científico es una ley puramente experimental, no habrá ninguna razón particular para escribirla en forma de una función compleja en lugar de en forma de un polinomio adecuado. Es conveniente seguir el uso científico denominando a las leyes puramente experimentales, es decir, leyes que son conocidas únicamente como resultado de una generalización a partir de observaciones y que todavía, en cualquier caso, no han sido deducidas de ninguna teoría bien comprobada, leyes "empíricas". Este es un uso más restringido del término "empírico" que el que de él hacen algunos filósofos que clasificarían incluso a las leyes teóricas como "empíricas", simplemente porque no son deducibles de proposiciones de la lógica y la matemática pura. Cuando un científico critica a una ley denominándola "meramente empírica" significa con ello mucho más que el que no sea deducible a partir de la lógica y la matemática pura, desde luego si lo que quería decir era esto el adverbio "meramente" resultaría del todo inapropiado.

A veces, aunque no estemos en posesión de una teoría que sugiera una determinada combinación de determinadas funciones matemáticas (tales como seno, logaritmo y funciones exponenciales) proporcionando una ley, hay buenas razones a menudo para ver si podemos ajustar nuestras conclusiones experimentales a una ley compuesta de tales funciones. Ello es así debido a que sabemos por experiencia que tales funciones aparecen con frecuencia en la física y en todas las demás partes. La función de seno está asociada, desde luego, con los fenómenos ondulatorios. Las leyes exponenciales aparecerán, evidentemente, cuando el índice de aumento o disminución de una determinada cantidad sea proporcional al valor de dicha cantidad. Por ejemplo, el índice de disminución del radio en una

roca es proporcional a la suma del radio de la roca, y el índice de aumento en una población de conejos estará en proporción al tamaño de la población, con tal que no aparezcan factores limitadores, tales como muerte por desnutrición o demasiados animales voraces. Por ello, a menudo contaremos con razones *cuasi*-teóricas para conjeturar una ley exponencial, por ejemplo, cuando un determinado polinomio se ajusta a nuestros datos, dentro de los límites del error experimental y observacional. En el caso de leyes que se refieran a tres o más variables, no es fácil encontrar un polinomio adecuado, y habrá que calcular, normalmente, funciones matemáticas más complicadas en cualquier caso.

Una característica de las leyes experimentales es su limitada exactitud. Por ejemplo, la ley de Boyle: "PV=constante" es verdadera únicamente en relación a un gas ideal; es decir, no es del todo verdadera con relación a ningún gas real. Algo más próximo a la verdad será la ecuación de Van der Waals: "$\left(P+\frac{a}{V^2}\right)(V-b)$=constante". Pero incluso ésta no es del todo verdadera con relación a gases reales. En realidad, la ley de Boyle se deduce de un modelo que supone que las moléculas de un gas son masas en forma de puntos. La ley de Van der Waals se acerca más a la verdad: se deduce de una ley que presupone que las moléculas no son masas en forma de puntos, sino esferas elásticas de volúmenes no insignificantes. Ninguno de estos modelos se corresponde con nuestro conocimiento teórico más depurado acerca de los gases (y es por ello por lo que los consideramos como modelos nada más y no como teorías).

La ley de Boyle y la ley de Van der Waals describen la variación de la presión con el volumen en un gas que está a una temperatura constante. Cuando tomamos en consideración la temperatura t, se convierten en "$PV=(t+c)$" y "$\left(P+\frac{a}{V^2}\right)(V-b)=R\,(t+c)$", respectivamente, siendo R y c constantes. Hagamos que $T=t+c$, y obtenemos la bien conocida expresión "RT" para el lado derecho de la ecuación. T es la temperatura absoluta. (Si t es temperatura en grados centígrados, entonces $c=273$, aproximadamente.) Puede darse ahora un paso que presenta considerable interés filosófico. En el momento en que se descubrió la ley de Boyle, se definieron la presión, P, y el volumen, V, como en la mecánica y en la geometría, respectivamente, pero la temperatura se definía en términos de graduaciones iguales de un termómetro de mercurio. ¿Por qué graduaciones *iguales*? En este momento existiría solamente la razón bastante empírica de que si hacemos los termómetros utilizando otros líquidos en lugar de mercurio nos encontramos con que las temperaturas tal como se definen mediante las graduaciones iguales de estos termómetros coinciden con las que se definen mediante las graduaciones iguales en un termómetro de mercurio, y así parecía como si la temperatura definida de este modo se correspondiese, aproximadamente al menos, con algo de tipo bastante sencillo de la naturaleza. El descubrimiento de la ley de Boyle

debió de haber reforzado esta sospecha, que más tarde vino a ser justificada por la teoría cinética, que demostraba que la temperatura de un gas, tal como aparece toscamente definida por los termómetros de mercurio, de hecho se aproxima mucho a algo de la naturaleza (en realidad a la energía cinética media de las moléculas de un gas). Sin embargo, el paso conceptual al que aludíamos al principio de este parágrafo no fue preciso que aguardase a la teoría cinética. El paso consiste en *definir* la temperatura mediante el comportamiento de un gas ideal (es decir, uno que obedezca exactamente a la ley de Boyle). El comportamiento de un gas ideal no puede ser descubierto examinando tales gases, puesto que no existen, sino que lo obtenemos mediante extrapolación a partir del comportamiento de los gases reales, que se aproximan, en grado diverso, a una conformidad exacta con la ley de Boyle. A partir de este momento "$PV=RT$" se utiliza para definir la temperatura, y el termómetro de mercurio viene a ser simplemente un indicador de la temperatura, al igual que el indicador de velocidad de un coche indica la velocidad. El termómetro ya no define la temperatura más de lo que el indicador de velocidad de un coche define la velocidad.

Puesto que la temperatura definida mediante "$PV=RT$" depende evidentemente de relaciones de la forma "PV=constante" no puede ser definida antes de que hayamos podido definir la presión y el volumen. Estos últimos son en sí mismos conceptos bastante complicados: la presión se define en términos de fuerza y área, siendo la fuerza un concepto teórico de la mecánica y el área (y el volumen) un concepto geométrico dependiente de la teoría geométrica. Es cierto que es posible dar una definición operacional burda del volumen (no dependiente de la teoría). Esto podría hacerse del modo siguiente: se dice que dos recipientes tienen el mismo volumen si una cantidad de líquido que llena exactamente uno, cuando se vierte en el otro, lo llenará exactamente también. Un recipiente A tiene un volumen igual a la suma de los volúmenes de los recipientes B y C si cuando se llenan ambos, B y C, con un líquido y este líquido se vierte en A, el líquido llena a A. Un sólido A tiene el mismo volumen que un recipiente B si la cantidad de líquido necesaria para llenar un recipiente C, en el que se ha colocado a A, es la misma que la necesaria para llenar C, si se pone dentro de C en lugar de A el líquido que llena a B. Es fácil ver que con esta base es posible definir el concepto de algo que posea $\frac{m}{n}$ veces el volumen de otra cosa, siendo m y n enteros. Es evidente, sin embargo, que un concepto tal de volumen, definido operacionalmente, es de aplicación muy limitada. Sólo podría aplicarse en situaciones relativas a objetos de tamaño mediano, fácilmente manejables: no podría aplicarse a cristales diminutos o a enormes objetos astronómicos, ni siquiera en contextos tales como la meteorología. En realidad, el análisis posterior sugiere que existen muy pocos campos en los que pudiera ser utilizado un concepto de volumen definido operacionalmente, para establecer ni siquiera leyes empíricas.

Con todo, si el volumen se *definiese* según el modo operacional que acaba de ser esbozado, constituiría lo que N. R. Campbell ha denominado magnitud *fundamental* [1]. Para que una propiedad [1] constituya una magnitud fundamental debe darse el caso de que se den las siguientes condiciones:

1) La propiedad puede ser poseída en distintos grados que posean un *orden*. Así pues, supóngase que se diese el caso de que la relación de *rayar* fuese asimétrica y transitiva, es decir: "(x) (y) x raya a $y \supset \sim y$ raya a x" y "(x) (y) (z) $[(x$ raya a $y \cdot y$ raya a $z) \supset x$ raya a $z]$" son ambas verdaderas. (Esto es casi, aunque no del todo, lo que ocurre en la realidad, ya que existen algunos casos excepcionales en los que una muestra de una sustancia rayará a una muestra de otra que rayará a una muestra de una tercera, y, con todo, la muestra de la tercera rayará la muestra de la primera.) Entonces la relación de *rayar* podría ser utilizada para definir la dureza como una propiedad, los grados de la cual poseerían un orden [2].

2) Existe una operación adecuada con las propiedades formales de la *adición*. Así, en el caso del volumen, el verter el líquido que llena exactamente el recipiente A junto con el líquido que llena exactamente el recipiente B en algún recipiente de mayor capacidad C que resulta entonces llenado exactamente, es una operación que corresponde formalmente a la operación aritmética de la adición, tal como se establece mediante la ecuación "$a+b=c$".

La dureza cumple, o casi cumple 1) pero no 2). El volumen definido operacionalmente según el modo acabado de mencionar cumple tanto 1) como 2) [3].

Campbell pasó luego a definir una magnitud *derivada* como sigue: Supóngase que X e Y son magnitudes fundamentales, y que en determinadas condiciones los valores de X e Y cumplen la ley empírica "$f(x, y)=c$", en donde c es una constante. Ahora háganse variar las condiciones alterando una determinada propiedad en los preparativos experimentales (por ejemplo, la temperatura). La ley "$f(x, y)=c$" nos permite definir la magnitud c de la propiedad que se ha variado. De este modo si (contrariamente a la realidad) la presión y el volumen fuesen magnitudes fundamentales, entonces la ley de los gases "$PV=RT$" definiría a la temperatura T como a una magnitud derivada. (Aquí R es, desde luego, simplemente una constante de proporción, de modo que permita unidades apropiadas para T, por ejemplo, hacer que la diferencia de temperatura

---

[1] O posiblemente una relación. Así, en una teoría relacional de la longitud (longitud en reposo, digamos, a fin de evitar complicaciones de la teoría de la relatividad), la magnitud *longitud* sería una relación, no una propiedad.

[2] Como en la escala de dureza de Mohs, que fue en un tiempo utilizada por los minerólogos. Esta estaba basada en el presupuesto de la asimetría y transitividad de la relación *raya* [2].

[3] En cuanto a (1): podemos decir que el volumen de un recipiente A es mayor que el del recipiente B si puede verterse líquido de A de modo que se llene por completo B, pero no puede verterse líquido de B de modo que llene por completo a A.

entre el hielo derretido y el agua en ebullición sea 100, como en la escala centígrada.) Y, desde luego, esto es lo que ha ocurrido con $PV=RT$, con las siguientes puntualizaciones: *a*) "$PV=RT$" es una idealización de las leyes empíricas acerca de los gases reales. *b*) La presión y el volumen no son magnitudes fundamentales en el sentido de Campbell, ya que son definidas en términos de los conceptos primitivos de la mecánica y la geometría, y estos conceptos primitivos no pueden ser definidos operacionalmente sino que es necesario que sean definidos implícitamente mediante el papel que desempeñan dentro de las teorías.

De hecho resulta dudoso que en el estado actual de la física exista ninguna magnitud fundamental importante en el sentido de Campbell, y por tanto nuestra explicación de leyes empíricas ha de tomar en cuenta el hecho de que los conceptos utilizados en ellas están cargados de teoría, aunque las propias leyes no tengan ninguna explicación dentro de las teorías que definen sus conceptos. Consideremos el concepto de longitud tal como resultaría si fuese definido operacionalmente como una magnitud fundamental en el sentido de Campbell y comparémoslo con el concepto de longitud tal como aparece realmente en la física.

Considérese un número de piezas de madera de forma recta. Definimos la rectitud operacionalmente sin hacer uso del concepto de longitud.

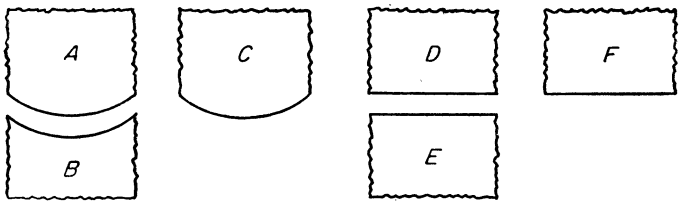

FIGURA 1.

En la figura 1, destinada a mostrar cómo las piezas de madera *A* y *B* encajan entre sí, *A* no encaja con *C* y *C* encaja con *B*. Decimos que los bordes de *A, B, C* no son rectos. En el caso de las piezas de madera *D, E, F, D* encaja con *E, E* encaja con *F* y *F* encaja con *D*. Decimos que las piezas de madera *son* rectas. También debemos presuponer que hemos establecido que las piezas de madera sean *rígidas,* siendo la rigidez definida operacionalmente como sigue: dos objetos son rígidos (o aproximadamente rígidos) si encajan entre sí (extremo con extremo, pongamos por caso) y no es fácil retorcerlos de modo que dejen de encajar mutuamente. Si dos piezas rectas rígidas de madera encajan entre sí, extremo con extremo, decimos que poseen la misma longitud. Si dos piezas rectas de madera se colocan extremo con extremo de modo que encajen con una tercera pieza recta de madera, decimos que la suma de las longitudes

de las dos primeras es igual a la longitud de la tercera. (La colocación de extremo con extremo proporciona aquí una operación física correspondiente a la adición.) Tomando un trozo de madera como medida patrón (como la vara métrica de París) determinamos una escala de longitud y construimos reglas calibradas. La longitud definida operacionalmente de este modo es una magnitud fundamental en el sentido de Campbell. Además, el concepto con ella relacionado de distancia entre dos objetos puede ser definido como la longitud de una pieza de madera recta que encajase extremo con extremo entre los dos objetos.

Debemos tomar nota, sin embargo, que tal concepto de longitud o distancia no es en modo alguno el concepto de longitud con que nos encontramos en física. Por lo pronto, el actual patrón de la longitud es la longitud ondulatoria de la luz de "cesium", que sólo puede ser comprendido mediante la teoría. Igualmente, cuando el físico habla acerca de la distancia de un electrón de su núcleo o del diámetro de una galaxia, evidentemente no está hablando acerca de lo que ocurre si colocamos objetos macroscópicos de tamaño mediano extremo con extremo. Podría no haber objeto macroscópico rígido alguno en el mundo y, sin embargo, las estrellas seguirían teniendo radios, los fotones seguirán poseyendo longitudes ondulatorias y las distancias entre las galaxias seguirían aumentando, sin duda. Desde luego, no habría habido seres humanos que averiguasen estas longitudes y distancias, e incluso habría sido imposible que las inteligencias puramente gaseosas pusieran en marcha la física experimental. Bien podría haberse dado el caso, como cuestión de *historia,* de que los hombres no hubiesen alcanzado el concepto científico, cargado de teoría, de longitud, excepto mediante un concepto más burdo al que hace justicia la explicación de Campbell de la medida fundamental. Considérese la primera vez en que la longitud ondulatoria de la luz de "cesium" (o cualquier otra) fue utilizada como patrón de longitud. Seguíamos necesitando escalas calibradas en nuestro espectrómetro y necesitábamos una retícula regulada, con el conocimiento del número de líneas que estaban espaciadas de modo equidistante en la retícula, y esto significa que necesitábamos utilizar un concepto de longitud más burdo, antes de que pudiésemos utilizar el más refinado. Es decir, necesitamos una medida fundamental, o algo parecido, antes de que podamos utilizar nuestros métodos más refinados. Este procedimiento puede ser justificado teóricamente, ya que existen razones teóricas para creer que las graduaciones que son equidistantes de acuerdo con nuestras convenciones de medida fundamental son aproximadamente equidistantes en términos de nuestro concepto teóricamente más refinado.

Vale la pena observar la diferencia entre un sistema propio de medida y el mero uso de instrumentos calibrados. El indicador de velocidad de un automóvil nos indica la velocidad del coche, pero no nos permite *medir* la velocidad. Es un *indicador* y no un *metro.* Lo que quiero decir con esto es que el indicador de velocidad (o su prototipo) es de esperar que esté calibrado mediante una serie de medidas de velocidad, que se

toman directamente midiendo la distancia viajada y el tiempo empleado en el viaje. La asimilación de la medida al registro en los instrumentos calibrados llevó posiblemente a Sir Arthur Eddington al muy desorientador enunciado de que las leyes científicas son acerca de *registros indicadores* (pointer readings) [3]. N. R. Campbell ha sugerido que la idea de Eddington de que la ciencia sea acerca de los registros indicadores puede ser una de las cosas que le hayan llevado a su filosofía idealista de la ciencia [4].

Puesto que las modernas ideas teóricas han venido a penetrar todo el aparato conceptual de la física, es difícil encontrar propiedad alguna que sea una magnitud fundamental en el sentido de Campbell. Las magnitudes, incluso en leyes completamente empíricas, pueden ser establecidas en términos de magnitudes teóricas. Considérese la ley de Ohm $\frac{E}{I}$=constante, donde E es la diferencia potencial entre los extremos de un conductor e I es la corriente a través del conductor. Es cierto que podría ser proporcionada una definición operacional de la suma de la corriente (por ejemplo, electroquímicamente, en términos de la suma de metal depositado por segundo en un electrodo en la solución de una sal determinada), pero es más natural considerar la corriente en términos teóricos, como la proporción en que los electrones pasan por un punto del conductor. Esta última definición percibimos que nos dice lo que la corriente eléctrica es realmente, mientras que la definición electroquímica proporciona simplemente un "*test*" para la existencia de la corriente de una determinada magnitud. Con todo, aunque los conceptos mediante los cuales se establece la ley pueden ser teóricos, la propia ley puede ser, sin embargo, puramente experimental, en el sentido de que no puede ser derivada de los axiomas de la teoría. Esto puede acontecer así por dos razons: *a*) La teoría quizá no sea lo suficientemente potente. *b*) La teoría puede ser lo suficientemente potente pero quizá no se hayan aplicado los suficientes mecanismos matemáticos al problema del descubrimiento de la derivación. En el caso de *a*) la presencia de leyes meramente empíricas es una muestra del estado insatisfactorio de conocimientos científicos. En el caso de *b*) no ocurre así, o al menos del mismo modo. Si la dependencia de leyes empíricas se debe simplemente a la incapacidad para tratar matemáticamente una situación muy compleja o para especificar con suficiente detalle la totalidad de un conjunto complejo de condiciones delimitadoras, no se debe, entonces, a la carencia de concepciones teóricas correctas de física. Por ejemplo, en el caso de los sucesivos intentos de conseguir por lo menos una aproximación remota a la ley de la conductividad de los electrolitos (que ha sido citada por Michael Scriven en un artículo [5] en el que mantiene el punto de vista de que una propiedad clave de las leyes físicas es la *inexactitud*), yo me inclino a suponer que las dificultades se deben a complejidades matemáticas y físicas y no a la carencia de leyes teóricas apropiadas o a inexactitud alguna en las leyes *básicas* de la física. En tal caso la incapacidad para explicar las leyes empíricas no

presenta mucho interés teórico: no muestra ningún punto débil en nuestra concepción de las leyes fundamentales de la naturaleza en la forma en que los experimentos que demostraban la no conservación de la paridad originaron un cambio en la teoría física básica.

Mientras que, en efecto, parece correcto el decir que las leyes empíricas pueden ser postuladas únicamente como aproximaciones, la afirmación de Michael Scriven [6] de que la inexactitud es una propiedad clave de las leyes físicas en general, es más dudosa. Además de las leyes empíricas este autor incluye casos tales como 1) las leyes del movimiento y la gravitación de Newton y 2) la primera y la segunda ley de la termodinámica. Estas son leyes teóricas, no empíricas, ya que no son generalizaciones directas a partir de observaciones, sino postulados de una teoría, y son comprobadas indirectamente comparando las consecuencias de la teoría como un todo con hechos experimentales u observacionales [4]. Podemos admitir que las leyes teóricas pueden ser consideradas como aproximaciones en aquellos casos en que proporcionan predicciones aproximadamente correctas, aun cuando las teorías a las que pertenecen han resultado ser falsas. Pero, contrariamente a la opinión de Scriven, estos casos *no* muestran que la mayoría de las leyes de la naturaleza sean inexactas, ni que la mayoría de las leyes que han sido consideradas leyes de la naturaleza no sean realmente leyes de la naturaleza en absoluto porque las teorías en cuestión hayan resultado ser incorrectas. No hay nada que nos impida afirmar que el progreso de la ciencia no ha demostrado, simplemente, que la mecánica newtoniana o la termodinámica clásica eran aproximaciones, sino que las ha *falsificado* [5].

En su réplica al comentario de Henryk Mehlberg [7], que cita la ley de que toda carga eléctrica es un múltiplo de una unidad única, y sugiere que, en general, las proposiciones acerca de cantidades discontinuas irán en contra de la tesis de Scriven, Scriven admite que algunas leyes son proposiciones "de las que no se conoce ninguna excepción" [8]. ¿Pero, no resulta desorientador el hablar de excepciones? El problema que presenta la ley de la gravitación de Newton no es el que posea unas cuantas excepciones: proporciona predicciones falsas siempre, aun cuando las dis-

---

[4] Existe, sin embargo, un sentido en el que puede probarse una ley individual. Supóngase que nos encontramos con que una teoría T no se ajusta del todo a los hechos observacionales, y que, en cambio, una teoría T', que se diferencia de T solamente en que una ley particular de T ha sido sustituida por otra, sí se ajusta a los hechos. Entonces podemos afirmar que la mayor parte de T está bien y que el problema radica, únicamente, en una de sus leyes. La situación es simplemente un caso especial de aquél en que todas o algunas leyes de T han de ser sustituidas, como en la sustitución de la mecánica newtoniana por la relatividad especial o general.

[5] La mecánica estadística muestra las incorrecciones de la segunda ley de la termodinámica. La primera ley todavía sigue en pie, en lo que se refiere a la propia mecánica estadística, pero supongo que se podría decir que resulta falsificada por la conversión de masa en energía y viceversa. El propio Scriven no alega que la *tercera* ley de la termodinámica sea falsa o inexacta.

Falsificado: es decir, que el progreso de la ciencia *ha demostrado que eran falsas. (Nota de la Traductora.)*

crepancias, que sabemos deben existir por deducción a partir de la teoría de Einstein, son en la mayoría de los casos demasiado pequeñas para que puedan ser detectadas mediante observación. A la vista de las muestras de leyes sin excepción, Scriven modificó su tesis, convirtiéndola en lo siguiente: no es muy *importante* el que las leyes de la naturaleza no tengan excepciones [9]. Cambiando "no tengan excepciones" por "no sean del todo exactas", yo estaría de acuerdo que esto es lo que ocurre en el caso de las leyes empíricas. Y puesto que las leyes son empíricas no existe ninguna razón teórica para preocuparse acerca del hecho de su ligera inexactitud. Sin embargo, en el caso de una ley teórica, las inexactitudes y excepciones son, por supuesto, muy importantes: incluso la menor discrepancia muestra que algo está mal en la teoría (a menos, desde luego, que podamos explicar la discrepancia —por ejemplo, cuando podemos mostrar que se debe a una forma inadecuada de llevar a cabo la experimentación—). Esto puede llevar a una modificación de la teoría, o incluso a la sustitución de la teoría por otra redicalmente opuesta (por ejemplo, una teoría ondulatoria sustituida por una teoría de partículas).

Las teorías consideradas como retratos (pictures) de la realidad

Si, de acuerdo con el punto de vista mantenido al final del parágrafo precedente, consideramos que una inexactitud dentro de una teoría es una objeción seria a la misma, de forma que se sigue que una teoría correcta *no* es inexacta, *ni* es, tampoco, una mera aproximación en el sentido en que lo es una ley empírica, se plantea el problema de si existe alguna razón para creer que *alguna* de nuestras teorías sea correcta. Esto forma parte del problema filosófico de la inducción, del que nos ocuparemos en el próximo capítulo, pero se origina de modo natural, aparte de derivarse de algunas sutiles consideraciones filosóficas, a partir de la reflexión de que en el pasado todas nuestras teorías físicas fundamentales han sido derrocadas a favor de otras mejores, y de que no existe razón alguna para creer que nuestras actuales teorías permanecerán en el futuro. Ahora bien, si decimos que todas las teorías en último análisis resultan ser falsas, ¿qué motivo existe para considerarlas como algo real, en lugar de considerarlas como meros instrumentos para proporcionar predicciones de utilidad práctica? Con toda seguridad, queremos seguir con la idea de que, aun cuando la teoría general de la relatividad resulte desplazada, seguirá siendo considerada como "más cercana a la verdad" que la teoría de la gravitación newtoniana. A menos que la ciencia avance mediante una especie de sucesiva aproximación a la verdad, ¿cuál es la ventaja *teórica* (en oposición a la ventaja *práctica*) de sustituir una teoría por otra? Hemos de admitir que si realizamos observaciones que hacen falsa a una teoría T, de modo que sustituimos T por T', entonces podemos decir que T ha sido falsificada por T', mientras que T' es todavía válida. Con todo, si echamos un vistazo a la historia de la ciencia, resulta bastante

probable que T' habrá de ser, a su vez, falsificada. ¿En qué sentido, pues, podemos decir que cuando T es sustituida por T' estamos más cerca de un retrato correcto del mundo? Supóngase que T es una teoría de partículas que resulta ser sustituida por una teoría de campo T', la cual será sustituida en el futuro por una teoría de partículas T'', la cual, a su vez, será sustituida por una teoría de campo T'''... Considérense ahora las dos siguientes hipótesis metafísicas: *a*) el mundo consiste en partículas, y hablar acerca de campos es sólo un modo de hablar acerca de las interacciones de las partículas, y *b*) el mundo consiste, en realidad, en un campo espacio temporal, y hablar acerca de partículas es, en realidad, únicamente hablar de los aspectos particulares del campo. Por lo que se refiere a la ontología de las teorías, T, T', T'', T''', se diría que estábamos oscilando fuertemente entre la verdad y la falsedad en la secuencia de teorías, ya sea *a*) o ya sea *b*) verdadera. Con ello no quiero sugerir que *a*) y b) agoten todas las posibles alternativas, como tampoco me atrevería a sugerir que "has dejado de pegar a tu mujer, ¿sí o no?" sea una forma correcta de preguntar. Supongamos únicamente que o bien *a*) o *b*) es verdadera en realidad y la otra es falsa.

En algunos casos, desde luego, resulta evidente lo que significaría afirmar que una teoría es mejor retrato del mundo que otra. Supóngase que el espacio fuese euclidiano. Entonces sería evidente que existe un sentido en el cual una geometría no euclidiana de curvatura constante $2R$ proporcionaría un retrato de la realidad mejor que una geometría no euclidiana semejante de curvatura constante R. En este caso las dos teorías no difieren ontológicamente sino cuantitativamente, tal como se refleja en la exactitud relativa de sus predicciones. Pero, ¿cómo compararíamos, pongamos por caso, una teoría que fuese notablemente correcta ontológicamente pero cuyas predicciones fuesen bastante inexactas, con una teoría que fuese incorrecta ontológicamente, pero que hiciese predicciones bastante exactas? A mi parecer, nos encontramos en este caso con un problema insoluble para aquellos filósofos que desean, como es mi caso, dar una explicación realista de las teorías científicas. No surge este problema con relación a aquellos filósofos de la ciencia que consideran las teorías físicas como meros instrumentos para la predicción de nuestras experiencias; en otras palabras, aquellos filósofos de la ciencia que pueden ser clasificados como fenomenalistas, operacionistas o instrumentalistas. Intentaré demostrar, sin embargo, que existen serias objeciones a tales filosofías de la ciencia no realistas.

A aquel que adopte una filosofía instrumentalista de la ciencia puede preguntársele *para* qué son instrumento las teorías. La respuesta dada generalmente es que son instrumentos para hacer predicciones; respuesta que debe ser formulada, si el punto de vista instrumentalista acerca de las teorías se mantiene en serio, en un lenguaje observacional no teórico, o neutral, con relación a las teorías. Las consideraciones de las págs. 63-8 sugieren que tal lenguaje observacional neutral con respecto a las teorías es difícil de encontrar. Normalmente las sentencias observacionales

están cargadas de teoría; por ejemplo, "la corriente de ese cable es de 5 miliamperios". Si la sentencia observacional fuese formulada de un modo neutral con relación a la teoría, por ejemplo, "la aguja de ese instrumento está enfrente del número 5", haría falta que fuese suplementada con reglas de correspondencia apropiadas, por ejemplo, "cuando la aguja de un instrumento [seguida de una descripción adecuadamente detallada del instrumento] señala el número '5' entonces la corriente que va a través de él es de 5 miliamperios". Tales reglas de correspondencia son más fáciles de esbozar que de establecer; en el ejemplo anterior, por ejemplo, he omitido el establecerlo de un modo probabilístico adecuado, ya que es seguro que ningún miliámmetro es siempre perfecto, y podría darse el caso de que la aguja del miliámmetro estuviese enfrente de "5" y, sin embargo, la corriente que pasase por el miliámmetro fuese distinta de 5 miliamperios.

Si mantemenos una filosofía realista de la ciencia no hay inconveniente alguno en manifestar que cuando miramos un miliámmetro podemos observar directamente que la corriente que pasa por él es de 5 miliamperios. El milámmetro funciona como una ampliación de nuestros sentidos ordinarios. Por supuesto, tenemos que aprender algo de teoría a fin de poder llevar a cabo tales observaciones: la percepción no es simplemente una percepción pasiva de las sensaciones sino que implica procesos cerebrales activos. Muchos de ellos son, quizá, innatos en nosotros, pero, desde luego, no somos conscientes de ellos. Considérese, por ejemplo, el modo en que las señales de los ojos permiten que el cerebro realice computaciones basadas en la convergencia de los ojos, proporcionándonos la habilidad para (de modo mecánico) juzgar las distancias de los objetos cercanos. En el caso del miliámmetro el proceso de información depende también de técnicas aprendidas de modo consciente. Si consideramos la cuestión desde este ángulo, la posibilidad de error en los miliámmetros no difiere epistemológicamente de la posibilidad de error de los sentidos humanos. Tal posibilidad de error no tiene por qué ser de resultados desastrosos, ya que al igual que las ilusiones de los sentidos ordinarios pueden ser descubiertas al realizar nuevas percepciones sensibles (también posiblemente falaces), así también los errores de los instrumentos científicos pueden ser descubiertos realizando nuevas observaciones científicas (también posiblemente falaces).

Puesto que los registros observacionales son ellos mismos indirectamente comprobables (por ejemplo, al hacer nuevas observaciones), no hay necesidad de suponer que la ciencia tiene que alzarse sobre una roca absolutamente infalible. No hay inconveniente en admitir que pudiera ser que estuviésemos equivocados en *cualquier parte* siempre que no tengamos que admitir que *probablemente* estemos equivocados en *todas partes*. El no haber apreciado este matiz es quizá una de las causas de la doctrina filosófica del *fenomenalismo,* según la cual todos los enunciados relativos a objetos materiales son enunciados acerca de datos sensibles reales y posibles (es decir, acerca de lo que los datos sensibles

son *si* se tienen otros datos sensibles). Los enunciados acerca de datos sensibles expresan únicamente lo que se ve, se oye, se siente, se huele, o se paladea, y así "yo tengo un dato sensible relativo a una rata rosa", puede ser verdad, evidentemente, con independencia de que la rata que esté ante nosotros sea rosa, o incluso aunque no haya ante nosotros rata alguna. Según el fenomenalismo la ciencia teórica se reduce a un sistema de reglas mediante el cual se pueden correlacionar y predecir nuestras experiencias sensibles. Un pionero de este tipo de filosofía de la ciencia (si bien la pureza de su fenomenalismo se vio infectada por las ramificaciones teológicas posiblemente poco importantes de su versión, y por ello tal versión difería en algunos aspectos del fenomenalismo moderno) fue el Obispo Berkeley [10]. Aproximadamente a comienzos de siglo fue introducida con gran impacto una filosofía fenomenalista por el filósofo austríaco de la ciencia Ernst Mach [11], y tal posición resultó del agrado de gran número de físicos, quienes, quizá, no la han sometido a una crítica filosófica demasiado rigurosa.

El instrumentalismo y el operacionismo

Puesto que este libro trata de la filosofía de la ciencia y no de la metafísica general, no entraré aquí en la discusión del fenomenalismo en la filosofía general, sino que indicaré simplemente que ha sido objeto de cierta crítica demoledora. En particular, remitiría al lector interesado al capítulo "Phenomenalism" de la obra de Wilfrid Sellars, *Science, Perception and Reality* [12]. Existen, sin embargo, determinadas filosofías de la ciencia, tales como el instrumentalismo y el operacionalismo, que son muy semejantes al fenomenalismo. Al igual que el fenomenalista mantiene que todos los enunciados relativos a objetos físicos (ya sean objetos comunes, como mesas, u objetos científicos, como electrones), son, en realidad, relativos a los datos sensibles, de modo que estas doctrinas reducen la ciencia teórica a una forma de hablar acerca de objetos macroscópicos. De este modo, según el punto de vista instrumentalista, las teorías no consisten en enunciados significativos acerca del mundo, si no que son mecanismos computacionales, que nos permiten predecir lo que veremos cuando miremos instrumentos de laboratorio y otros objetos de tamaño medio. El operacionalismo difiere del instrumentalismo en mantener que las teorías científicas consisten en enunciados significativos, pero comparte la inclinación del instrumentalismo a negar la interpretación realista de los enunciados teóricos, es decir, la inclinación a negar que son relativos a entidades teóricas con existencia independiente (a saber, electrones, fotones, etc.), Según el operacionalismo, todos los conceptos de la física teórica tienen que ser definidos en términos de operaciones con objetos de tamaño medio tales como instrumentos de laboratorio. Según este punto de vista la relación entre los enunciados acerca de electrones y los enunciados acerca de placas fotográficas, por ejemplo, es la misma

que la relación entre enunciados acerca de placas fotográficas y enunciados relativos a datos sensibles en la posición fenomenalista.

La doctrina del operacionalismo fue defendida por el físico P. W. Bridgman, aunque, al propio tiempo, le disgustaba el mundo grandilocuente del "operacionalismo" [13]. Según él, todos los conceptos de la ciencia debían ser definidos en términos de las operaciones que deben ser realizadas a fin de determinar si el concepto puede ser aplicado. En el caso de los conceptos cuantitativos, deben ser definidos en términos de operaciones de medida. Así, el "electrón" debía ser definido mediante referencia a operaciones con cámaras oscuras Wilson o con células fotoeléctricas. La "longitud" debía ser definida mediante operaciones con reglas. (Recuérdense, sin embargo, las objeciones a esto vistas anteriormente en la página 102.) Todo esto resulta un poco vago, y todavía se hace más vago cuando Bridgman concede que los conceptos teóricos son introducidos, en parte al menos, mediante operaciones con "papel y lápiz". (Considérese un concepto tan abstruso como el giro isotópico en la mecánica cuántica, por ejemplo. Es evidente que tal concepto no puede relacionarse directamente con operaciones experimentales. Si puede ser relacionado con experimentos es sólo después de una cantidad considerable de trabajo teórico.) Admitir operaciones con papel y lápiz significa, desde luego, trivializar la doctrina del operacionalismo, ya que parece ser que cualquier tipo de definición es válida, incluso algún tipo de definición implícita en términos de postulados de teoría. Lo defendido por el operacionalismo parece reducirse a una vaga insistencia en la necesidad de que las teorías científicas sean comprobables de modo experimental u observacional, en algún sentido. Y, de ser así, la cuestión puede ser, con toda seguridad, formulada en un lenguaje más preciso y sin la desorientadora alusión a "operaciones". Este tipo de alusión es especialmente desorientador en cuanto sugiere, de forma completamente gratuita, que cuando hablamos de electrones, neutrones, etc., estamos hablando, en realidad, acerca de instrumentos de laboratorio y de marcas de lápiz sobre trozos de papel.

Bridgman ha apoyado su defensa del operacionalismo alegando que fue una actitud operacionalista la que llevó a Einstein a la teoría *especial* de la relatividad [14]. En el mismo escrito Bridgman critica la teoría *general* de la relatividad porque le parece que no se debe a una actitud operacionalista. Yo sugeriría, por el contrario, que el enfoque de apariencia operacionalista que encontramos en el escrito original de Einstein acerca de la teoría especial de la relatividad [15] no era esencial al escrito y que le mermaba claridad. La definición de Einstein de la simultaneidad, en términos de observadores con varas de medir y relojes, es del todo artificial, evidentemente. Podemos discutir si dos acontecimientos aquí y en "Arcturus" son simultáneos o no con referencia a un conjunto de coordenadas sin ninguna mitología relativa a un observador de "Arcturus" equipado con un reloj y capaz de remitirnos señales luminosas. Si una teoría necesita hablar de esto, ello sería una prueba de debilidad de la teoría, ya que no tenemos noticias de ningún observador tal en "Arctu-

rus". Lo esencial de la definición de la simultaneidad de Einstein puede ser dado, de hecho, sin la mitología operacionalista, y esto es lo que se hace con frecuencia en los modernos tratados de la materia [16]. Yo me atrevería a conjeturar que la forma operacionalista en que formula Einstein la relatividad especial se debió, quizá, a un positivismo filosófico, que prendió en él de forma bastante superficial, y que más tarde llegó a rechazar [6]. Desde luego, se podría argüir que aun cuando el operacionalismo no es esencial para la teoría especial de la relatividad puede haber sido de valor heurístico al permitir que Einstein llegase a esta teoría. Tal argumentación es irrelevante para la *verdad* del operacionalismo; pues, incluso un cuento de hadas podría poner casualmente la idea correcta en la mente de un científico. A lo que se podría añadir la objeción de que el pensamiento *no* operacionalista también ha sido de valor, como, por ejemplo, en el desarrollo de la teoría *general* de la relatividad de Einstein. Es verdad que para Bridgman la teoría general es insatisfactoria a causa de la dificultad para poder dar una explicación operacionalista de ella, pero la objeción que aquí hace, a no ser que se caiga en un círculo vicioso, debe estar basada en razones filosóficas tales que puedan ser ofrecidas a favor del operacionalismo, independientemente del apoyo que pueda proporcionar la actuación real de Einstein u otros científicos.

Si el operacionismo, ampliado con la inclusión de "operaciones con papel y lápiz", ha de ser algo de mayor interés que una simple insistencia en que las teorías han de ser comprobadas, debe ser interpretado, creo yo, como una forma del instrumentalismo. El instrumentalismo sí permite un libre juego de construcciones teóricas, pero según él las teorías son meros formalismos sin interpretar [7] mediante los cuales podemos predecir el comportamiento de objetos macroscópicos, por ejemplo, instrumentos de laboratorio. El instrumentalismo puede ser criticado en dos sentidos: *a*) desde un punto de vista metafísico; *b*) desde un punto de vista metodológico. Comenzaré por la crítica metafísica.

---

[6] El que Einstein simpatizase en algún tiempo con el positivismo aparece con referencia a Hume y Mach en su autobiografía intelectual, en la cual también muestra claramente que en períodos posteriores de su vida, por lo menos, llegó a rechazar de modo completamente explícito, cualquier tendencia positivista [17].

[7] A tales formalismos se les denomina, a veces, "algoritmos" [18]. Esto puede ser desorientador, ya que en la lógica y en la matemática "algoritmo" normalmente significa "procedimiento de decisión", y es fácil deducir del teorema de Church que no existe ningún algoritmo de decisión para un cálculo lo suficientemente rico como para expresar una teoría física. Es decir, en la filosofía de la ciencia "algoritmo" se usa a veces de modo distinto a como se usa en la lógica matemática, lo cual parece lamentable. Por la misma razón puede ser desorientador decir que las partes matemáticas de la física son "mecanismos computadores", ya que esto podría sugerir falsamente que existen procedimientos de decisión y que no es necesario ingenio alguno para demostrar teoremas en la física matemática.

Crítica metafísica del instrumentalismo y el operacionismo

El atractivo del instrumentalismo radica, en parte, en el hecho de que conocemos las entidades submicroscópicas solamente mediante observaciones realizadas en las macroscópicas: contemplamos placas fotográficas, amperímetros, papel tornasol en pasta, y así sucesivamente. Una interpretación de las teorías científicas que sugiere que éstas son únicamente ayudas para hablar acerca del comportamiento de objetos macroscópicos se diría que economiza esfuerzos por cuanto nos ahorra el tener que postular entidades cuya existencia es quizá dudosa. Este punto de vista aparece frecuentemente reforzado por teorías positivistas del significado. Puesto que no podemos señalar, u observar directamente, las entidades microscópicas de la física, y se considera que el significado puede ser conferido a las palabras únicamente mediante el acto de señalar objetos perceptibles u operaciones experimentales, parecería, pues, que es cuestionable el que las teorías físicas posean significado alguno, y quizá no debieran ser tratadas sino como meros formalismos no interpretados, excepto en cuanto se conectan con observaciones u objetos macroscópicos. Si tal doctrina mantiene que lo que pueda decirse de un modo significativo que existe debe ser, *en principio,* perceptible, se originan dificultades. ¿Como podríamos decidir si es *lógicamente* posible o no percibir un electrón o un neutrino? Si se dice que es *físicamente* imposible ver o tocar un electrón o un neutrino, desde luego esto es del todo cierto, pero los realistas añadirían que esto es así, simplemente, a causa de las leyes de la naturaleza relativas a los electrones y a los neutrinos.

Incluso en una teoría muy empirista del significado, la afirmación de que las sentencias que son ostensiblemente relativas a objetos demasiado pequeños para ser percibidos carecen de significado, es incoherente. Supóngase que se admite (lo cual yo, desde luego, no admitiría) que los términos científicos tienen significado si y sólo si su significado puede ser dado ostensiblemente (o, en parte, ostensiblemente) del mismo modo en que, quizá, puedan ser introducidas palabras como "mesa", "cuadrado" y "rojo". Asimismo, podemos suponer que la palabra "percibir" se introduce de un modo tal, también, ya que ¿de qué otro modo podría ser establecida la tesis relativa al significado de "en principio perceptible"? Entonces la descripción "partículas menores que ninguna de las que fuesen, sean o serán percibidas" o (si la palabra modal "pueda" está permitida en la teoría empirista del significado) "partículas menores que ninguna que pueda ser percibida" debe ser significativa de acuerdo con el criterio empirista del significado.

La posición instrumentalista se formula a veces diciendo que las teorías físicas no son afirmaciones acerca del mundo sino "billetes de inferencia" (véase Gilbert Ryle [19]), o "técnicas de explicación" (véase

Stephen Toulmin [8]). Sea o no sea esto cierto, las teorías científicas no se presentan, ciertamente, de modo tal que parezcan reglas de inferencia. Se presentan como un sistema de proposiciones que pueden ser deducidas a partir de un conjunto de postulados, que son las leyes fundamentales de la teoría. No existe ninguna regla eficaz para decidir cuáles son los teoremas de una teoría, excepto, posiblemente, en el caso de teorías muy débiles. Cualquier teoría medianamente potente tendría que ser formalizada dentro del cálculo de predicados completo, y es consecuencia del teorema de Church el que no exista ningún método eficaz (procedimiento de decisión) para separar los teoremas de este cálculo. Sólo podemos saber que algo es un teorema si encontramos el modo de derivarlo a partir de los postulados. Con todo, quizá el instrumentalista siga manteniendo que los postulados de una teoría son reglas, y que las reglas de inferencia son reglas de *segundo orden* para deducir reglas derivadas. Es interesante observar que al tratar de destacar el carácter regulativo, más que proposicional, de una teoría, Toulmin hace uso del ejemplo de la óptica geométrica, que no necesita ser afirmado proposicionalmente, sino que podría ser considerado como una técnica para realizar diagramas.

El instrumentalista no necesita decir, como el operacionalista tendría que hacer, que las sentencias teóricas son equivalentes o traducibles a determinadas clases de sentencias relativas a objetos macroscópicos observables. El instrumentalista permite que los conceptos teóricos sean construcciones del teórico, libres por completo. Sin embargo, el instrumentalista estaría de acuerdo con el operacionalista en mantener que en la ciencia ningún enunciado se formula acerca de entidades que no sean observables macroscópicas. El operacionalista mantendría esto basándose en que los enunciados teóricos son traducibles a enunciados acerca de objetos observables, mientras que el instrumentalista negaría la traducibilidad. De acuerdo con el instrumentalista las sentencias teóricas no expresan, en realidad, enunciados acerca de entidades debido a que, en realidad, no expresan enunciados en absoluto: expresan reglas para realizar predicciones acerca de objetos observables. Puesto que tanto el operacionalista como el instrumentalista están de acuerdo en lo esencial, es decir, en que las teorías científicas se ocupan, en realidad, de objetos observables macroscópicos, la diferencia entre ellos no es importante, metafísicamente. Por esta razón me propongo reunir a las dos teorías juntas, como variedades de lo que podría denominarse "macro-fenomenalismo", para distinguirlo del fenomenalismo metafísico ordinario que reduce todo a datos sensibles [9].

[8] En su capítulo "Theories and Maps", en *The Philosophy of Science* [20]. La comparación con mapas sugiere, en principio, un punto de vista realista respecto a las teorías, pero, conforme Toulmin desarrolla la analogía, la idea parece consistir principalmente en que las teorías nos capacitan para encontrar nuestro camino con relación a los *fenómenos,* al igual que los mapas nos capacitan para encontrar nuestro camino cuando vamos al campo. No está del todo claro que la posición de Toulmin sea realmente instrumentalista, pero la interpretación más lógica de lo que él afirma, a mi parecer, es instrumentalista.

[9] Un fenomenalista también será macro-fenomenalista, pero no hace falta que

Evidentemente, el modo más eficaz de defender el macro-fenomenalismo sería dando una prueba de que los términos teóricos se pueden eliminar siempre de la ciencia. Antes de que tal prueba, o intento de prueba, puede tener lugar es necesario, evidentemente, dividir el vocabulario científico en dos clases: términos observacionales y términos teóricos. Ya hemos visto motivos para dudar de que pueda hacerse esto. La especificación de *sentencias* observacionales mediante su vocabulario (términos observacionales únicamente) varía de la explicación pragmática de las sentencias observacionales, según la cual una sentencia puede ser considerada sentencia observacional con tal que en un contexto perceptual particular un científico pueda decidir rápidamente acerca de su verdad o falsedad sin ningún pensamiento o cálculo explícito [10]. De acuerdo con la explicación pragmática de las sentencias observacionales no hay razón para que estas sentencias no contengan términos teóricos, es decir, términos que no serían inteligibles al sentido común no científico. De este modo "esta placa fotográfica muestra una huella de partícula de beta curva" o incluso "la corriente de ese cable es de diez miliamperios" pueden ser sentencias observacionales, si la explicación pragmática de las sentencias observacionales es correcta. De ser así, la caracterización de las sentencias observacionales mediante referencia al vocabulario es poco prometedora. A pesar de todo, sigamos con el macro-fenomenalista de momento y admitamos con él, por seguir la discusión, que existe una diferencia practicable entre los términos observacionales y teóricos de una teoría. Hagamos, también, que los términos lógicos, tales como "algún", "no", "y", así como también los términos de la matemática pura (incluyendo la sintaxis) sean considerados como parte del vocabulario observacional. Esto parece razonable, ya que la lógica y la matemática pura pueden ser comprendidas por una persona que no conoce física alguna. Podemos, pues, dividir el vocabulario V de una teoría científica en dos partes, V_0 y V_T, respectivamente, de las cuales V_0 contiene los términos observacionales, lógicos y matemáticos, y V_T contiene los restantes. Consideraremos ahora dos mecanismos que pueden ser propuestos a fin de demostrar que V_T es dispensable.

El primero de ellos es la *sentencia de Ramsey* [21]. Sea A un conjunto de postulados de una teoría científica. Supóngase que el conjunto de A es la sentencia

$$\text{"}F(\Phi_1, \Phi_2, \ldots, \Phi_n)\text{"}, \text{ en donde } \text{"}\Phi_1\text{"}, \text{"}\Phi_2\text{"}, \ldots, \text{"}\Phi_n\text{"}$$

son los predicados que aparecen en A y que pertenecen a V_T. Podemos escribir "$F(\Phi_1, \Phi_2, \ldots, \Phi_n)$" en notación teórica de conjuntos como "$F'(K_1, K_2, \ldots, K_n)$", en la que cada K_i es una clase de m-múltiplos de los cuales el predicado Φ_i m-ésimo es verdadero. Sustitúyase ahora "$F'(K_1, K_2, \ldots,$

un macro-fenomenalista sea fenomenalista; es decir, puede ser realista con relación a objetos observables de tamaño medio, tales como sus instrumentos de laboratorio.

[10] Véanse las págs. 66-67 anteriores y las referencias a Feyerabend y Carnap.

K_n)" por la sentencia "$(\exists X_1), (\exists X_2), \ldots, (\exists X_n)\ F'(X_1, X_2, \ldots, X_n)$" en la cual las constantes "K_1", "K_2", ..., "K_n" han desaparecido en favor de variables ligadas. Esta es la sentencia de Ramsey de la teoría. Por ejemplo, si "Φ_k" es el predicado "es un electrón", entonces "K_k" será la expresión "clase de electrones", y en la sentencia de Ramsey tenemos justamente "existe un X_k tal que..." [11].

La sentencia de Ramsey nos permite, desde luego, prescindir del *vocabulario* V_T, pero el defensor de una filosofía realista de la ciencia puede seguir negando que esta maniobra tenga alguna implicación ontológica. La posibilidad de sustituir los axiomas de una teoría de electrones que contengan el predicado "... es un electrón" por un adecuado "$(\exists K_k)$ $(\ldots \epsilon K_k \ldots)$" no impide que este último enunciado haga una afirmación existencial equivalente a la de la existencia de electrones. El presupuesto de la verdad de la sentencia de Ramsey pone límites a la pertenencia a las clases K_i, parecidos, precisamente, a aquel impuesto por la teoría original a las extensiones de los predicados "es un electrón", etc. La afirmación "existen electrones", por ejemplo, es sustituida entonces por la parte de la sentencia de Ramsey "$\ldots (\exists K_k) \ldots (\exists x)\ (\ldots x \epsilon K_k \ldots)$" que afirma que existe una clase no vacía K_k. Presupongamos también que "... es demasiado pequeño para ser percibido" y "... es capaz de producir huellas en cámaras oscuras" forman parte del vocabulario observacional [12]. Entonces si la teoría original implica

"(x) (x es un electrón $\supset x$ es demasiado pequeño para ser percibido y x es capaz de producir huellas en cámaras oscuras)"

entonces del mismo modo la sentencia de Ramsey implica

"$\ldots (\exists K_k) \ldots (\exists x) \ldots (x \epsilon K_k \supset x$ es demasiado pequeño para ser percibido y x es capaz de producir huellas en cámaras oscuras").

Lo que hace en realidad la técnica de la sentencia de Ramsey es sustituir una teoría acerca de electrones y protones (pongamos por caso) por una teoría acerca de determinadas clases de cosas que son especificadas

---

[11] Es más habitual escribir la sentencia de Ramsey simplemente en la forma

$$(\exists \phi_1)(\exists \phi_2) \ldots (\exists \phi_n) F(\phi_1, \phi_2, \ldots \phi_n)$$

cuantificando los predicados, en lugar de las clases. Sin embargo, yo sigo a Quine al experimentar grandes dudas con relación a que sea adecuado cuantificar los predicados [22], y, por tanto, he preferido expresar la sentencia de Ramsey como una cuantificación de clases.

[12] No me gusta gran cosa incluir las expresiones modales "ser percibido" y "capaz de producir" como parte del vocabulario observacional, o de otro tipo, de una teoría, pero me resulta difícil inventar un ejemplo adecuado sin ellas. Yo mismo consideraría esto como una severa crítica a la idea de definir las sentencias observacionales mediante referencia a un vocabulario, idea que he aceptado a los defensores del sistema de la sentencia de Ramsey tan sólo con objeto de proseguir la discusión.

de modo abstracto, de tal modo que las relaciones entre estas cosas sean isomórficas con aquellas relaciones que la teoría original mantenía que se daban entre electrones y protones. Si la teoría original afirma la existencia de electrones, la sentencia de Ramsey contendrá una parte de la forma "... $(\exists K_k)$ ... $(\exists x)\, x \epsilon K_k$" y, por tanto, en realidad supondrá la misma implicación ontológica.

Es cierto, desde luego, que *tanto* la teoría original *como* su sentencia de Ramsey equivalente *podría* (por lo que atañe a la anterior argumentación) ser interpretadas *ambas* instrumentalmente y sin implicaciones ontológicas. Pero de modo igual *ambas,* la teoría original y su sentencia de Ramsey equivalente, podrían ser interpretadas realísticamente. El mecanismo de la sentencia de Ramsey no corta el hielo metafísico ni de un lado ni de otro.

Otro método para eliminar la parte V_T de una teoría procede del teorema de Craig [23]. Para comprender lo que es esto será conveniente una breve explicación preliminar. Cuando uno piensa en un sistema axiomático quizá uno piensa en un conjunto de axiomas *finito.* Sin embargo, a veces uno encuentra útil, en la lógica y la matemática, el construir sistemas con conjuntos de axiomas infinitos. De este modo en una axiomatización finita de la lógica sentencial podemos especificar "$p \supset (q \supset p)$" como axioma, y puede haber una *regla de sustitución,* una regla de inferencia, que autoriza la sustitución de letras de sentencias por fórmulas bien formadas. Evidentemente, puede lograrse lo mismo sin contar con ninguna regla de sustitución, pero especificando un conjunto infinito de axiomas como sigue: cualquier fórmula de la forma "$A \supset (B \supset A)$" es un teorema. Se sobreentiende que los lugares de "A" y "B" en la fórmula anterior tienen que ser ocupados por fórmulas bien formadas de la lógica sentencial y, puesto que existe un procedimiento efectivo (o recursivo) para decidir si una fórmula de la lógica sentencial está o no está bien formada, no hay dificultad para decidir si cualquier sentencia dada pertenece o no al conjunto de axiomas infinito en cuestión [13]. (Los sistemas de axiomas infinitos pueden, desde luego, resultar particularmente convenientes dentro de la lógica, en cuanto que nos permiten prescindir de la regla de sustitución.)

En general, un conjunto de axiomas infinito es permisible si existe un modo efectivo de decidir si una determinada sentencia pertenece o no al conjunto. Es decir, debe haber un *procedimiento de decisión* o algoritmo que nos permita mostrar que *a*) la sentencia está en el conjunto, si está en él, y *b*) la sentencia no está en el conjunto, si no está en él. Por ejemplo, no valdría especificar como conjunto de axiomas infinito a todas las fórmulas válidas del cálculo de predicados. Esto se debe a

[13] Desde luego es evidente que restringirse a fórmulas bien formadas de la *lógica sentencial* es innecesario, y que un conjunto infinito de axiomas "$A \supset (B \supset A)$" puede ser utilizado como parte de la axiomatización de la lógica de predicados, teoría elemental de números o teoría de conjuntos, o cualquier otra teoría para la que pueda proponerse una especificación efectiva de lo que es una fórmula bien formada.

que, de acuerdo con el teorema de Church, no habría método efectivo alguno para decidir si una fórmula era un axioma o no. Desde luego, existe un método para mostrar (en última instancia) que una fórmula es una fórmula válida de la lógica de predicados si es que lo es, pero nunca podemos decir mediante un método mecánico, que una fórmula *no* sea una fórmula válida de la lógica de predicados. Por tanto, por mucho tiempo que llevemos aplicando nuestro mecanismo para sacar fórmulas válidas, si la fórmula que se está considerando no ha aparecido, nunca podemos determinar si ello se debe a que no hemos sacado fórmulas válidas durante tiempo suficiente para que apareciese o a que no aparecería *nunca* por mucho que siguiésemos sacando fórmulas válidas. Igualmente, todavía valdría menos el especificar como conjunto de axiomas a todas las verdades de la teoría elemental de números. Como corolario del teorema de la incompletitud (incompleteness) de Gödel, no existe ningún mecanismo para sacar todas las verdades de la teoría elemental de números, y mucho menos mecanismo alguno para sacar todas las no verdades. Sin embargo, un conjunto de axiomas infinito es correcto siempre que exista un procedimiento efectivo para decidir si una sentencia pertenece al conjunto o no. *A grosso modo* podemos decir que un conjunto de axiomas infinito es correcto siempre que sea posible [14] construir un computador que nos proporcione una respuesta de "sí" o "no" a una pregunta de la forma: "¿Está o no está la sentencia S en el conjunto?

Ahora sea T una teoría física determinada cuyo vocabulario es V, conteniendo las partes V_0 y V_T. El teorema de Craig muestra cómo podemos sustituir T por una teoría T' que no contenga ninguno de los términos de V_T, y en la cual todos los teoremas de T que contienen únicamente expresiones de V_0 son también demostrables. Es decir, a pesar de la falta de aparato conceptual de V_T la teoría T' es tan potente, por lo que se refiere a la prueba de teoremas en V_0, como lo era la teoría original T. Con todo, el vocabulario de T' es más amplio que el de V_0, puesto que complementamos V_0 con un aparato metalingüístico que nos permite hablar acerca de las expresiones de la teoría original T y expresar la sintaxis de T. De este modo "A es demostrable en T", será expresable en T'. Esto se hace mediante el recurso de la numeración de Gödel, con lo cual se asocia un número natural único a cualquier secuencia de expresiones o secuencia de secuencias de expresiones del vocabulario de T. "A es demostrable en T" es traducida así a una sentencia de la teoría elemental de números, y "demostrable en T" corresponde a

---

[14] "Posible", aquí *por lo menos* significa "matemáticamente posible". Pero en la práctica se necesita un sentido más fuerte de "posible". ¡Evidentemente de nada serviría que hubiese un algoritmo para decidir la cuestión de si una sentencia estaba en un conjunto de axiomas o no, pero que el algoritmo fuera tan lento en su funcionamiento que incluso con sentencias de complejidad moderadas a un computador eléctrico le llevase años conseguir la respuesta! Un conjunto infinito de axiomas, tal como se vislumbra en el teorema de Craig, no cumpliría, desde luego, este requisito práctico.

una función aritmética recursiva, es decir, una que nos permite determinar, de un modo mecánico, si cierta propiedad aritmética es válida respecto a un número o no. En particular, podemos determinar si es válida respecto al número Gödel asociado con alguna sentencia *A*.

Ahora sea *A* un teorema de *T* que contiene únicamente los términos V_0. (Por supuesto, una prueba de *A* dentro de *T* contendrá, con toda probabilidad, términos de V_T.) Estipulamos ahora que la sentencia $A \cdot A \cdot A \cdot \ldots A$ (una conjunción de *n* *A*s) será axioma de *T'* si y sólo si *n* es el número Gödel de una prueba de *A* en *T*. Es fácil ver que es correcto especificar un conjunto infinito de axiomas de este modo, es decir, que existe un medio efectivo de determinar si una sentencia dada de *T'* está en el conjunto o no. Ya que si contamos el número de *A*s de $A \cdot A \cdot A \cdot \ldots A$ (habrá montones de ellas[15]) podemos averiguar si este número es el número Gödel de una secuencia de secuencias de expresiones de *T* y, además, si este número Gödel es el de una secuencia de secuencias de expresiones que constituye una prueba de *A* dentro de *T*.

Es evidente que el "quid" de la cuestión radica en sustituir *T* por una teoría *T'* que no contenga parte alguna del vocabulario de *T* pero que nos proporcione el modo de hablar acerca de las *expresiones* teóricas de *T*. Si *T* contiene la palabra "electrón", entonces *T'* contendrá la palabra "'electrón'". Podría pensarse entonces, que la posibilidad de sustituir *T* por *T'* sugiere una interpretación instrumentalista de *T*. Si podemos arreglárnoslas hablando acerca del *lenguaje* de *T*, por ejemplo, acerca de la palabra "electrón", en vez de hablar acerca de electrones, ¿no sugiere esto que cuando parece que hablemos acerca de electrones, en realidad solamente estamos utilizando un mecanismo elaborado de inferencia que nos permite pasar de determinadas sentencias V_0 a otras determinadas sentencias V_0? Por el contrario, yo sugeriría, sin embargo, que el realista podría replicar, con bastante más razón, que las consideraciones mencionadas no muestran nada por el estilo. Podría señalar que el método de Craig sólo puede ser aplicado una vez que hayamos construido la teoría *T*, y que el éxito de *T'* se explica por el hecho de que la teoría original *T* es verdadera respecto de las cosas de las que trata ostensiblemente, en otras palabras, por el hecho de que realmente existan electrones o aquello que sea postulado por la teoría *T*. De no existir tales cosas, y si *T* no fuese verdadera de un modo realista, ¿no sería del todo inexplicable el éxito de *T'*? Habría que suponer que existían innumerables accidentes afortunados relativos al comportamiento de las cosas mencionadas en el vocabulario observacional, de modo tal que se comportaban milagrosamente *como si* fuesen producidas por las cosas no existentes de las que se hablaba ostensiblemente en el vocabulario teórico.

Esto nos lleva a un aspecto general relativo al instrumentalismo. Un instrumentalista es, desde luego, instrumentalista con relación a todas las teorías, mientras que un realista puede ser realista con relación a

[15] Esta consideración va en contra de la posibilidad *práctica* (en oposición a la matemática) de utilizar una teoría como *T'*. Véase nota 14, anteriormente.

algunas teorías (aquellas que él cree que son verdaderas) e instrumentalista con relación a otras (las que él cree que son útiles, pero no verdaderas). La diferencia entre los dos tipos de teoría queda señalada en el lenguaje científico denominando al contenido del último tipo de teoría "modelo". Desde luego una seria objeción que se puede hacer al instrumentalismo es la de que olvida esta importante distinción [24]. La principal objeción realista al instrumentalismo es, como ya he sugerido, la de que desde el punto de vista instrumentalista es un hecho sorprendente e inexplicable el que el mundo sea tal que los hechos V_0 puedan estar relacionados con hechos V_0 mediante sentencias V_T, mientras que desde un punto de vista realista no hay nada misterioso en ello: si las sentencias V_T se refieren realmente a los mecanismos ocultos del universo, después de todo, entonces puede esperarse, *desde luego,* que nos ayuden a relacionar sentencias V_0 con otras sentencias V_0. Considérese un hombre (en el siglo dieciséis) que sea realista con relación a la hipótesis copernicana pero instrumentalista con relación a la de Ptolomeo. Este hombre puede explicar la utilidad instrumental del sistema ptolomeico de los epiciclos ya que puede demostrar que el sistema ptolomeico puede proporcionar casi las mismas predicciones relativas a los movimientos aparentes de los planetas que la hipótesis copernicana. De aquí que el presupuesto de la verdad realista de la hipótesis copernicana explique la utilidad instrumental de la ptolomeica. Tal explicación de la utilidad instrumental de determinadas teorías no sería posible si *todas* las teorías fuesen consideradas como meramente instrumentales. De modo semejante, es difícil creer que las regularidades que aparecen a nivel macroscópico fuesen de tal modo como *si* se debiesen a cosas del nivel sub-microscópico. (Al igual que en el fenomenalismo ordinario existe el interrogante de por qué los datos sensibles tales como los del rabo de un gato en la parte izquierda de un sofá deben ir seguidos de los datos sensibles de la cabeza y los bigotes de un gato a la derecha del sofá. Desde el presupuesto realista de que nuestras experiencias sensibles son causadas por un gato con existencia independiente que camina detrás del sofá, tengamos o no tengamos datos sensibles algunos, esta secuencia de datos sensibles se explica inmediatamente, por supuesto.) Si no tenemos más que leyes instrumentales éstas pueden proporcionar explicaciones en el sentido de que nos permiten predecir, pero no pueden proporcionar explicaciones en el sentido de explicar los datos en bruto.

Debe añadirse que quien acepte el uso instrumentalista del teorema de Craig tendrá que, *qua* descubridor científico, fingirse, ante sí mismo, realista en el contexto del descubrimiento. Solamente a teorías acabadas puede ser aplicado el método de Craig. Sin una actitud al menos en algún sentido realista no es probable que se pudiera imaginar la teoría, en primer lugar. Hempel ha señalado [25] que puesto que el método de Craig requiere una sistematización puramente deductiva no nos permite razonar inductivamente, como en el siguiente ejemplo: A causa de los datos observacionales O_1, O_2, ..., O_n, una determinada sustancia es, probable-

mente, del tipo teórico S_i (es decir "S_i" es parte de V_T) y, así, pues, probablemente produce el efecto S_k (aquí "S_k" es también parte de V_T), ya que nuestra teoría afirma que si algo es del tipo S_i conduce a algo del tipo S_k. Cualquier cosa del tipo S_k lleva al resultado observacional O_m. Así que probablemente O_m. Una nueva dificultad para el macro-fenomenalista es la siguiente: éste mantiene que las sentencias ostensiblemente relativas a electrones son, simplemente, meros instrumentos de cálculo, y, por tanto, puede decirse que los electrones son ficciones teóricas. Pero si los electrones son ficciones teóricas y las plumas estilográficas son objetos reales, ¿dónde trazamos la línea entre lo real y lo ficticio? ¿Son ficción las bacterias? (Sólo pueden ser observadas al microscopio.) ¿Lo son los virus? Estos no pueden ser vistos con un microscopio ordinario, sino con un microscopio electrónico. ¿Lo son las moléculas de las proteínas? (que no son mucho menores que los virus). ¿Es una ficción teórica todo lo que no puede ser visto por el ojo desnudo? Si es así, entonces las cosas que sólo pueden verse a través de un cristal de aumento son ficciones, y, sin embargo, se parecen mucho a las cosas que pueden ser vistas por el ojo desnudo.

Debe observarse que la objeción relativa a "¿dónde trazamos la línea?" es una objeción al macro-fenomenalista, que es realista con respecto a los objetos físicos de tamaño medio, pero no es una objeción al fenomenalista filosófico ordinario, que mantiene que el mundo está constituido por datos sensibles. Para él, tanto los electrones como las mesas navegan en el mismo barco: ambas cosas son ficciones teóricas. Mas yo doy por descontado que el fenomenalismo ortodoxo no es una posición filosófica sostenible. Una cuestión particularmente espinosa para el macro-fenomenalista es ésta: Supóngase que dice que una molécula es una ficción teórica, mientras que una mota de polvo visible no lo es. ¿Qué dirá con relación a un cristal pequeño pero visible? Grover Maxwell ha señalado que, de acuerdo con la teoría de valencias moderna, un cristal puede ser considerado como una sola macro-molécula, en cuyo caso algunas moléculas son, con toda seguridad, no entidades teóricas sino observables [26].

En cualquier caso la noción de "cosa observable" dista mucho de ser una cuestión clara. ¿Hemos de decir que no observamos mesas sino únicamente zonas coloreadas de luz y sombra? ¿No es una interpretación de los datos de la experiencia visual el que lo que vemos es algo sobre lo que podemos escribir y no una pieza camuflada de tramoya hecha de lona? Si descartamos aquellas interpretaciones que el sentido común hace conforme con la información que es transmitida en pulsaciones electroquímicas a través de los nervios ópticos, tendremos que decir que ni siquiera las mesas son observables y volveremos a la teoría del fenomenalismo ortodoxo[16]. Si la percepción ordinaria depende inevitablemente de nuestro es-

---

[16] Bien pudiera alegarse que los "datos sensibles" del fenomenalismo ortodoxo son entidades teóricas más dudosas que lo son los electrones. A menos que "Yo tengo un dato sensible de mesa" sea sólo una forma peculiar de decir "me parece

quema conceptual ordinario, ¿por qué no puede estar formada la percepción del científico, parcialmente, mediante su conocimiento teórico? Por ejemplo, si ve un chispazo en una pantalla ¿no puede decir "vi a un electrón dar en la pantalla"? Es cierto que el chispazo pudiera despistarle en este caso, y quizá se equivoque al interpretar su experiencia visual como la de un electrón dando en la pantalla, pero, exactamente igual, las percepciones del sentido común están expuestas a engaño, lo cual los "magos" saben aprovechar muy bien. De modo semejante, al contemplar un amperímetro, ¿por qué no puede decir el científico que observa una corriente de diez amperios, en lugar de la coincidencia de un indicador con un número? Todo esto sugiere la teoría pragmática de las sentencias observacionales, a la que hemos aludido anteriormente en la pág. 67. Como ya hemos indicado, ésta ha sido defendida con vigor por P. K. Feyerabend [27], y fue en un tiempo defendida por Rudolf Carnap [28].

Feyerabend, de forma inspirada, considera al sistema perceptual humano (ojos y cerebro) como un instrumento científico. Los instrumentos científicos, tales como termómetros y galvanómetros, constituyen simplemente una extensión del cuerpo humano, y, dada la adecuada preparación científica, podemos utilizarlos para afirmar *inmediatamente* la existencia de cosas tales como temperaturas y corrientes eléctricas, exactamente igual que, supuesto simplemente el conocimiento vulgar, y como resultado de fotones chocando contra la retina, un hombre puede afirmar *inmediatamente* que existe un objeto tridimensional, duro, de madera, delante de él. Cuando consideramos las observaciones desde este ángulo no vemos razón alguna para trazar una línea tajante entre el "lenguaje de la observación vulgar" y el lenguaje de la teoría científica. (Cuando digo que el azúcar de mi taza de té se ha *disuelto,* ¿estoy hablando en sentido vulgar o científico?) Resulta imposible hacer una distinción entre un lenguaje teórico y un lenguaje observacional para demarcar una línea entre las "cosas reales" y las "ficciones teóricas". Podemos, desde luego, considerar ficciones a las cosas mencionadas ostensiblemente en teorías falsas, pero sin embargo útiles en la práctica. Pero en el caso de teorías que han superado las pruebas de comprobación y que (en la actualidad, por lo menos) se consideran verdaderas, no tenemos motivos para negarnos a tratar las entidades allí mencionadas como objetos de observación verdadera.

Crítica metodológica del instrumentalismo

Popper ha argumentado que una teoría instrumentalista de la física es incapaz de explicar el interés del científico por comprobar implicaciones remotas de su teoría [29]. Si una teoría fuese simplemente un mecanismo práctico de cálculo, entonces el fallo ante una prueba de comprobación

---

que hay una mesa", pero una charla de datos sensibles tan adulterada no es lo suficientemente fuerte para soportar el peso teórico del fenomenalismo.

no supondría una refutación de la teoría sino simplemente una limitación a su aplicabilidad. Esta crítica *metodológica* al instrumentalismo ha sido desarrollada con acierto por Feyerabend. En contraste con Popper, Feyerabend mantiene que la mecánica cuántica *de hoy en día* sí que exige una interpretación instrumentalista [30]. Sin embargo, Feyerabend mantiene que los científicos no deberían dejar la cuestión así, sino que deberían buscar activamente una teoría que fuese capaz de una interpretación realista. Puesto que una teoría instrumentalista es mucho menos refutable que una teoría realista, el instrumentalismo lleva a un contentarse con las pruebas de una teoría. Feyerabend considera la colisión que tuvo lugar en el siglo dieciséis entre las hipótesis ptolomeicas y copernicanas acerca de los movimientos planetarios. Si Copérnico, Galileo y otros, hubieran aceptado todos ellos el punto de vista instrumentalista, como quizá lo hizo el Cardenal Bellarmine, oponente de Galileo, entonces no habría sido aceptada la teoría copernicana. En la fase en la que se encontraba la ciencia, la hipótesis ptolomeica era más conveniente como instrumento de predicción, en parte porque era la técnica aceptada de aquel momento, y por ello podía ser aplicada más fácilmente. Dentro de los límites de la observación que eran entonces posibles, podía hacerse fácilmente (con la postulación de epiciclos suficientes) que proporcionase predicciones relativas a la posición de los planetas, al menos con tanta exactitud como pudiera hacerlo la hipótesis copernicana. También podía ser adaptada fácilmente a observaciones incluso más exactas simplemente añadiendo más epiciclos. Según Feyerabend era éste su principal problema: al acomodarse ella misma a observaciones cada vez más exactas mediante la técnica de construir epiciclos extras se protegía a sí misma contra posible refutación, y se hizo incomprobable. De este modo se convirtió en una barrera al progreso científico, que sólo puede darse a través de teorías que están expuestas a la refutación.

Digamos que una teoría es comprobable en sentido débil si observaciones contrarias pudieran llevar únicamente a una modificación bastante simple de ella, pero que es comprobable en sentido fuerte solamente si la teoría no siempre puede ser salvada mediante una modificación tan sencilla, *ad hoc*. Así pues, si se hubiese dado el caso de que la hipótesis ptolomeica pudiera haber sido salvada siempre mediante la adición de epiciclos apropiados, entonces la teoría habría sido comprobable débilmente, pero no fuertemente comprobable. Es característico de las teorías instrumentalistas el ser débilmente comprobables, únicamente. Por este motivo Feyerabend mantiene que convendría que los descubrimientos científicos y las innovaciones diesen como resultado la sustitución de la mecánica cuántica contemporánea por una teoría fuertemente comprobable. Pues, según él, la moderna mecánica cuántica *no* es susceptible de una interpretación realista, sino que requiere la interpretación instrumentalista asociada con la escuela de Copenhague que la hace débilmente comprobable únicamente. Es cierto que la interpretación de Copenhague ha sido postulada por sus defensores de un modo en el que se utilizan malos

argumentos filosóficos. Algunos de estos argumentos son semejantes a los del fenomenalismo filosófico ordinario, y serían argumentos en contra de cualquier tipo de teoría realista. Otros dependen de la inferencia a partir de "no podemos averiguar si existe o no tal cosa" de "no existen tales cosas en absoluto", que es un *non-sequitur*. Se dice a menudo, por ejemplo, que la interacción entre un instrumento de medición y un electrón nos impide ser capaces de descubrir a la vez la posición e impulso de un electrón por debajo de los límites de exactitud impuestos por el principio de indeterminación de Heisenberg. Se concluye entonces que el electrón no posee posición ni impulso definidos. Esto o bien es un *non-sequitur* o, de lo contrario, depende de presupuestos positivistas muy discutibles, que figuran como premisas implícitas. Con todo, mantiene Feyerabend, la interpretación de Copenhague puede ser defendida mediante buenos argumentos *físicos* que hay que reconocer aun cuando hayamos rechazado los malos argumentos *filosóficos*.

La mecánica cuántica y el instrumentalismo

Discutamos una de las consideraciones físicas que hacen surgir dificultades al considerar a un electrón como una partícula real, y que, consecuentemente, sugieren una filosofía instrumentalista de la mecánica cuántica. Se trata del famoso experimento de las "dos ranuras". Véase la figura 2. Considérese un punto de origen *S* que emite electrones. Estos pasan a través de dos ranuras *A* y *B* en una pantalla *X*. Los electrones

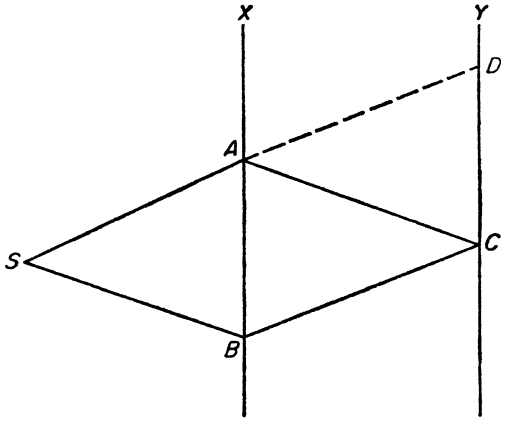

Figura 2.

dan, entonces, en una nueva pantalla *Y*, donde su impacto es observado como un oscurecimiento de una superficie fotográfica o como chispas en una superficie fluorescente. (En un experimento real las "ranuras" son

excisiones en una reja de cristal, y la situación es ligeramente más complicada, pero la presente explicación idealizada es suficiente para lo que nos proponemos.)

Si ambas ranuras están abiertas la corriente de electrones producirá una serie de bandas de interferencia (regiones oscuras sobre una placa fotográfica) en la superficie *Y*, con su mayor intensidad en el punto central *C* y con intensidades que disminuirán gradualmente conforme nos movamos a las bandas situadas por encima o por debajo de *C*. El fenómeno es exactamente lo que esperaríamos si la corriente de electrones consistiese en realidad en ondas. Fenómenos de interferencia semejantes son predictibles en la teoría clásica de la óptica física. Hasta aquí el experimento sugiere que deberíamos considerar a los electrones no como partículas sino como ondas.

Supóngase ahora que la intensidad del rayo del electrón que sale de *S* se reduce. Cuando se hace muy bébil lo que se observa no es que las bandas de interferencia en *Y* se hagan simplemente muy difusas. Lo que se observa es que los electrones individuales chocan con la pantalla *Y* en diversos lugares. Si *Y* es una superficie fluorescente, el impacto que puede observarse pueden ser diminutos chispazos. Se verá que los chispazos aparecen en las regiones en donde aparecían antes las bandas de interferencia, y por tanto estas bandas representan regiones donde es elevada la probabilidad de impacto de los electrones. Estos hechos sugieren que deberíamos abandonar la imagen ondulatoria del electrón en favor de una imagen de partículas.

La imagen de partículas no servirá, sin embargo, ya que lo que se necesita para el patrón de interferencia es la existencia de *ambas* ranuras. Esto es inexplicable en la imagen de las partículas, ya que, con toda seguridad, una partícula tiene que pasar por una ranura u otra. Si pasa a través de *A* es difícil ver qué puede importar el que la ranura *B* esté abierta o cerrada. Con todo, sí importa. Si *B* está cerrada entonces el lugar más probable en que un electrón dé en *Y* no es *C* sino *D*. Ahora bien, ¿cómo diablos sabe el electrón, cuando pasa a través de *A*, si *B* está abierta o cerrada? O, dicho de un modo menos antropomórfico, ¿qué influencia física puede ejercer una ranura abierta o cerrada en *B* sobre una partícula en *A*?

Estas dificultades no preocupan al adepto a la interpretación de Copenhague, ya que éste considera la teoría como simplemente un instrumento conveniente con el que hacer predicciones. Compagina, perfectamente, por tanto, el utilizar una imagen ondulatoria los lunes, miércoles y viernes (para predecir patrones de interferencia), y una imagen de partículas los martes, jueves y sábados (para predecir chispazos individuales). Aun cuando las dos imágenes son incompatibles entre sí, el uso de ambas (en momentos diferentes) permite que se hagan las predicciones correctas [31].

Debe observarse que, de acuerdo con la interpretación de Copenhague, el electrón no tiene simultáneamente ambas cosas, una posición definida y

un impulso definido. Cuando choca con *Y* su posición se hace definida, pero su impulso es del todo indefinido. Al pasar la pantalla *X* tanto su impulso como su posición son indefinidos, y por ello no debe decirse que ha pasado ya bien a través de la ranura *A* o a través de la ranura *B*. Consideremos ahora algunas de las dificultades que se originan si tratamos de proponer un alternativa a la interpretación de Copenhague y considerar, contrariamente a ella, al electrón puramente como una partícula con una trayectoria definida, como Popper ha intentado hacer [32]. Debe recordarse que la interpretación habitual de la función ondulatoria en la mecánica cuántica es la de que representa probabilidades. De este modo las bandas de interferencia de la pantalla representan regiones donde las ondas de probabilidad de *A* y *B* se refuerzan mutuamente. Estas probabilidades se interpretan habitualmente como frecuencias. Decir que la probabilidad de que un electrón dé en una pantalla en un lugar determinado es mayor que en otro, es decir que, por término medio, una mayor proporción de electrones dan en la pantalla en un lugar que en el otro. La primera parte de la explicación de Popper consiste en sustituir la concepción de frecuencia de la probabilidad por la interpretación de la tendencia. Esto ya lo hemos discutido en el capítulo 2, anteriormente (véanse págs. 42-44). Las tendencias son para Popper supuestos hipotéticos de la ciencia, al igual que lo son los electrones o las fuerzas. No hay dificultad en considerar a las tendencias como estados reales de entidades reales. Entonces, de acuerdo con Popper, un electrón en la situación del experimento de las dos ranuras, con ambas ranuras abiertas, tiene una tendencia real a ir a determinados lugares, tales como *C*, en la pantalla *Y*, más que a otros. Si alteramos la situación experimental de forma que sólo la ranura *A* esté abierta, entonces la tendencia cambia, de forma que la mayor parte de los electrones van a *D* y casi nunca van a *C*. De acuerdo con Popper, el electrón, en realidad, va, en efecto, a través de una ranura u otra, pero la tendencia en el caso de que *B* esté abierta es distinta de la tendencia cuando la ranura *B* está cerrada. Debemos recordar que, según Popper, una tendencia pertenece no solamente a un elemento de una situación experimental, tal como el electrón, sino a la situación globalmente. Si hablamos de la tendencia de un electrón a hacer algo, esto debe ser entendido como una elipsis para hablar de la tendencia del electrón *en una cierta situación* a hacer algo.

Con todo, todavía parece haber algo de misterio. La influencia de la ranura *B* en la alteración de la tendencia parece tener resonancias de acción a distancia, y es difícil ver con qué principios físicos podría ser explicada. Según la concepción de Popper, el electrón tiene tanto una posición definida como un impulso definido en todo momento (aun cuando el principio de Heisenberg nos impida determinar ambas cosas), y por tanto tiene una trayectoria bien definida. Se sigue de esto que si un electrón con un determinado impulso en la ranura *A* va a *C* mientras que otro va a *C'*, entonces el segundo electrón debe haber adquirido un impulso en la dirección *CC'* que el primero no poseía. Véase la figura 3. Es evi-

dente, por tanto, que si trabajamos con trayectorias bien definidas, el principio de conservación de impulso ha de ser abandonado.

Feyerabend indica [33] que en una comunicación privada Popper ha establecido que en su opinión las leyes de conservación aplican estadísticamente a grandes números de partículas, pero que no son válidas en los casos particulares. Feyerabend admite que es *posible* que la microfísica pudiera ser desarrollada de tal modo que las leyes de conservación (para casos particulares) fuesen rechazadas. Sin embargo, todavía nadie sabe cómo llevar esto a cabo, y en tanto que no se desarrolle tal teoría con una capacidad predictiva igual o superior a la de la mecánica cuántica actual, la solución de Popper al problema de dar una interpretación realista a la mecánica cuántica debe considerarse una especulación únicamente. Las leyes de conservación están bien atrincheradas en la física contemporánea y cuentan con mucho apoyo experimental. Feyerabend, desde luego, admite que el hecho de que cuenten con mucho apoyo experimental no es decisivo, ya que los experimentos han de ser interpretados a la luz

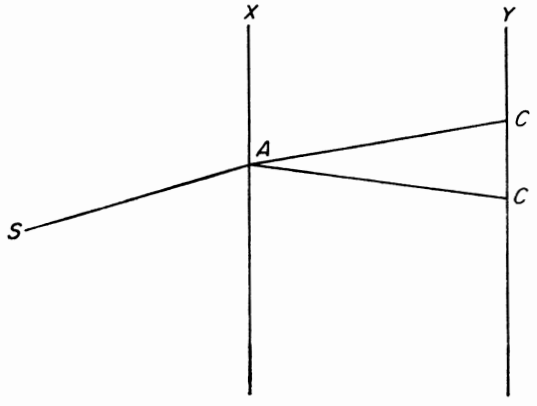

FIGURA 3.

de las teorías, y el apoyo que proporcionan a una teoría puede deberse al ámbito restringido en el que hemos intentado aplicar y comprobar nuestra teoría. (La teoría del movimiento aristotélica, que requiere que una fuerza mantenga a un cuerpo a velocidad constante, estaba bien respaldada por la experiencia cotidiana.) Con todo, es, cuando menos, muy dudoso el que pueda desarrollarse una teoría acertada que rechace las leyes de conservación. La posición de Feyerabend, por consiguiente, es la de que, atendiendo a razones *físicas,* es difícil no aceptar una actitud instrumentalista respecto a la mecánica cuántica pero que, sin embargo, por razones metodológicas, esto no es adecuado, y deberíamos poner empeño en la búsqueda de una teoría realista, y por tanto más radicalmente comprobable, de la mecánica cuántica.

Aun cuando le concedo a Feyerabend que existen muy buenas razones

físicas para aceptar una actitud instrumentalista hacia la mecánica cuántica (aunque los argumentos *filosóficos* a favor del instrumentalismo sean malos) quisiera sugerir que los argumentos físicos a favor del instrumentalismo no son *compulsivos*. Incluso en ausencia de una teoría mejor, y *más claramente* realista de la mecánica cuántica, ¿no podemos ser realistas acerca de la mecánica cuántica existente? El experimento de las dos ranuras (y otras consideraciones similares) con toda seguridad muestran únicamente que los electrones y semejantes son entidades bastante extrañas, por ejemplo, que no están situados con precisión en el espacio y que sus características son bastante distintas de aquéllas que nos son familiares en la física clásica. De este modo se diferencian en gran medida de los objetos familiares en que no poseen trayectorias e impulsos bien definidos. Además, algunas de las leyes a las que obedecen son de un tipo desacostumbrado. Margenau ha señalado [34] que el principio de exclusión de Pauli impide que dos electrones libres posean *exactamente* la misma velocidad, aun cuando se encuentren a billones de millas de distancia. (Sin embargo, la diferencia entre sus velocidades permitidas se hace tan pequeña que resulta insignificante excepto a distancias aproximadamente subatómicas.) Tal ley parece implicar acción a distancia. En esto, así como el no poseer trayectorias o impulsos bien definidos, los electrones son objetos extraños. Pero, ¿es necesario que nos neguemos a admitir que ellos son, al igual que otras partículas incluso más extrañas, reales? Los electrones no pueden ser caracterizables como ondas o como partículas, aunque la teoría puede considerar conveniente el tratarnos como cualquiera de estas cosas, según la ocasión [17]. De modo alternativo, podríamos mantener que la física moderna es instrumentalista, pero que puede constituir, sin embargo, una aproximación o una sugerencia de una teoría realista. Es decir, hay una teoría realista incipiente, aunque es todavía desconocida por nosotros. Ya que, de no ser así, ¿cómo podría explicarse el éxito de una teoría instrumentalista? De otro modo, el éxito de la teoría instrumentalista tendría que depender de muchas meras coincidencias a nivel macroscópico.

Leyes, teorías y lógica

Cualquier teoría matemática axiomatizable puede ser expresada en la terminología de la lógica de predicados añadiendo al vocabulario de la lógica de predicados determinados predicados constantes [18], e introduciendo determinados axiomas que impliquen estos predicados constantes. Consideraré que esto es así respecto a cualquier teoría de la física, aun cuando en la realidad no sea tarea sencilla axiomatizar una teoría de

[17] Margenau, en su libro *Open Vistas* [35] cita a A. N. Whitehead, quien mantiene que las cosas pueden existir sin localización simple.

[18] Por ejemplo, la teoría de conjuntos introduce el predicado binario constante "ϵ". Véanse págs. 33-35 anteriormente.

este modo, incluso en casos relativamente manejables, como se podrá ver con claridad, examinando la axiomatización del caso relativamente sencillo de la mecánica de partículas clásicas de McKinsey, Sugar y Suppes [36]. (E incluso aquí, los autores mencionados no formalizan la lógica subyacente sino que hacen uso de una teoría de conjuntos intuitiva.) Con una teoría física viva, tal como la mecánica cuántica, que está continuamente cambiando, y algunas de cuyas partes quizá no sean, incluso, del todo compatibles con otras, la axiomatización es difícilmente practicable. Con todo, es útil pensar en alguna futura axiomatización de cualquier fase histórica particular de la física, como un ideal teórico. Una ley de deducción L, a partir de otras más fundamentales B_1, B_2, ..., B_n, puede considerarse válida si la ley $(B_1 \cdot B_2 \cdot \ldots B_n) \supset L$ es deducible de la axiomatización ficticia. Lo que ocurre en la práctica en la física es que la proposición $(B_1, B_2, \ldots, B_n) \supset L$ se deduce utilizando medios que el lógico y el matemático puro no considerarían del todo rigurosos.

He estado presuponiendo aquí que las leyes de la naturaleza pueden ser formuladas dentro de una lógica extensional [19]. De este modo, si limitamos nuestra atención, en aras de la simplicidad, a una forma de ley sencilla, "todos los *A*s son *B*s", esto puede expresarse simplemente como "$(x)\, Ax \supset Bx$". En contra de esto, algunos filósofos mantienen que la ley anterior tiene que ser expresada por medio del denominado condicional fuerte, que puede ser simbolizado mediante "$\rightarrow$". De este modo la sentencia "$(x)\, Ax \supset Bx$" dice exactamente que nada es A sin ser B, y es vacuamente verdadera cuando, de hecho, nada es A. Por otra parte, la proposición "$(x)\, Ax \rightarrow Bx$" se dice que afirma que el que una cosa sea A hace que ésta *necesite* ser B, y en particular la no existencia de *A*s no asegura la verdad de "$(x)\, Ax \rightarrow Bx$". Por ejemplo, sea "A" el predicado "no es accionado por una fuerza" y "B" el predicado "se mueve con velocidad uniforme en línea recta". Entonces si la primera ley del movimiento de Newton es expresada mediante la fórmula "$(x)\, Ax \supset Bx$" se convierte en vacuamente verdadera si no hay cuerpos en el universo que no sean accionados por una fuerza, mientras que "$(x)\, Ax \rightarrow Bx$" no sería vacua en este sentido.

Sin embargo, otros filósofos, entre los cuales me encuentro, siguen manteniendo que una lógica extensional, en la que se utilice el condicional débil "$\supset$", es perfectamente adecuada para fines científicos. Ahora bien, si estos argumentos se pueden mantener, entonces, aunque sólo sea por razones de simplicidad, debemos evitar el condicional fuerte. También debe recordarse que Quine [37] ha formulado formidables argumentaciones en contra de la lógica modal cuantificacional, que es lo que se necesitaría para formalizar una teoría científica utilizando el condicional fuerte. Pero aunque pueda encontrarse una lógica modal cuantificacional que se defienda frente a las objeciones de Quine, como han postulado Rudolf Carnap [38] y otros, seguirá siendo bastante complicada y desordenada,

[19] Véanse págs. 37-38.

comparada con la simplicidad y elegancia de la lógica de predicados ordinaria, y por ello debemos considerar con atención los credenciales de cualquier argumento que se proponga demostrar la indispensabilidad de los condicionales fuertes en la ciencia.

Consideremos, antes de nada, la objeción de que una lógica extensional hace que algunas leyes sean vacuamente verdaderas. De este modo, probablemente la primera ley del movimiento de Newton es vacuamente verdadera, ya que en ninguna parte del universo existe cuerpo alguno no accionado por una fuerza. Una respuesta a esto sería: "¿Y qué?" Es decir, aunque la ley fuese vacuamente verdadera ¿tiene eso alguna importancia? Aun cuando sea *vacuamente* verdadera seguirá siendo *verdadera.* En cualquier caso el ejemplo de la primera ley del movimiento de Newton es bastante especial, ya que no parece ser necesitada en la mecánica newtoniana, ya que es un caso especial de la *segunda* ley, que establece que la fuerza es proporcional a la aceleración. Por tanto, evidentemente, si no existe fuerza alguna entonces existe cero aceleración.

Se invoca a veces la primera ley para definir un sistema inerte, pero, en realidad, no es necesaria. Considérese cómo puede definirse un sistema inerte en la mecánica celeste (la mecánica del sistema solar). Tenemos un número de cuerpos M_1, M_2, ..., M_n. Nos encontramos con que pueden elegirse los ejes, de tal modo que las aceleraciones mutuas de los cuerpos en cualquier instante determinado puedan ser descompuestas en sus componentes f_{12}, f_{21}, f_{13}, f_{31}, etc., de modo que f_{kl} y f_{lk} pertenezcan al k-avo y l-avo cuerpo, respectivamente, y estén en dirección opuesta, en la línea recta que une a estos cuerpos, y pueden ser asignados números reales m_1, m_2, ..., m_n a estos cuerpos, de forma que $m_k f_{kl} = m_l f_{lk}$. Tal sistema de ejes se denomina sistema inerte. Puede demostrarse que tal sistema de ejes debe existir. (Obsérvese, sin embargo, que esto *no* constituye una solución analítica al problema del cuerpo-n. Los valores reales de m_1, m_2, ..., m_n se determinan mediante un método de aproximación sucesiva.) Así es como se definen en astronomía los sistemas de ejes inertes, y no hay modo mejor de hacerlo. La ficción de un cuerpo que no es accionado por una fuerza no ayuda en absoluto en la práctica[20].

Yo sugeriría, pues, que si una teoría científica es formulada correcta-

[20] En una conversación mi amigo Richard Routley ha planteado la interesante cuestión de si será necesaria en la *relatividad especial* la primera ley de Newton a fin de definir un sistema inerte. En realidad, en la relatividad especial las relaciones simples de la tercera ley del movimiento de Newton ($m_k f_{kl} = m_l f_{lk}$) no tinen vigencia excepto en las colisiones, y el análisis es considerablemente más complicado. Con todo, a mí no me parece del todo claro que sea imposible, ni siquiera en la relatividad especial, dar una definición de "sistema inerte" sin hacer uso de la ficción de cuerpos que no son accionados por una fuerza. En la práctica, en la relatividad especial consideramos que un sistema inerte newtoniano se aproxima a un sistema de ejes inerte. El utilizar una entidad *ficticia,* un cuerpo no accionado por una fuerza, no lleva a ninguna parte, ni siquiera como aproximación. En *teoría,* en la relatividad especial podríamos apelar a un conjunto de ejes para los cuales la tercera ley fuese obedecida en fenómenos de colisión, pero el método sería, por supuesto, de poca utilidad en la práctica, ya que las colisiones son por fortuna escasas dentro de la mecánica celeste.

mente, no hay necesidad de que ninguno de sus axiomas resulte ser vacuamente verdadero. Con todo, *si* fuese necesario incluir leyes vacuamente verdaderas en el desarrollo axiomático de una teoría, tales leyes no pueden hacer daño alguno. Considérese una ley "$(x)\, Ax \supset Bx$" a la que le sucede ser vacuamente verdadera ya que no existe ninguna A. Nunca podríamos utilizarla para argumentar (supongamos) a favor de una conclusión falsa "Ba". Nuestra argumentación tendría la forma

1) $(x)\, Ax \supset Bx$
2) Aa
3) $\therefore Ba$

No podemos argumentar de este modo simplemente porque si 1) es vacuamente verdadera entonces 2) debe ser falsa, y no podemos, por ello, permitirnos el utilizarla como premisa. Resumiendo, el admitir verdades vacuas en un sistema no nos permite deducir ninguna falsedad. (Esto resulta, desde luego, evidente, si reflexionamos en que las verdades *vacuas* siguen siendo *verdades* vacuas.)

Supóngase que nunca se hubiese tomado la decisión, en ningún lugar del universo, de fabricar la bomba atómica. En este caso, la proposición

4) (x) (y) [x es un pedazo de uranio 235 (de tamaño mayor a uno determinado) e y es un pedazo de uranio 235 (del mismo tamaño) y x es impulsado contra $y \supset$ existe una colosal explosión]

habría sido vacuamente verdadera. Más aún, los científicos seguirían deseando afirmar la condicional *subjuntiva*

5) (x) (y) [si x *fuese* un pedazo de uranio 235 (de tamaño mayor a uno determinado) e y fuese un pedazo de uranio 235 (del mismo tamaño) y x fuese impulsado contra y entonces *habría* una colosal explosión]

y desearían negar

6) (x) (y) [si x *fuese* un pedazo de uranio 235 (de tamaño mayor a uno determinado) e y fuese un pedazo de uranio 235 (del mismo tamaño) y x fuese impulsado contra y entonces el Sol *se pondría* púrpura brillante].

Es decir, podría decirse, los científicos desearían afirmar la condicional subjuntiva 5) y negar la condicional subjuntiva 6). Ninguna de ellas podría ser vacuamente verdadera, y podría decirse que ambas tendrían que ser analizadas en términos del condicional fuerte. Sin embargo, aun cuando esto sea así, podemos replicar que un científico no se verá seriamente

afectado si le impedimos el uso de los condicionales fuertes o subjuntivos. En lugar de afirmar 5) puede expresar lo que piense afirmando la proposición *metalingüística*

7) La sentencia 4) se sigue de una teoría bien comprobada.

De modo semejante, en vez de negar 6), puede expresarse mediante la proposición metalingüística

8) La sentencia 9) no se sigue de una teoría bien comprobada, siendo 9) la sentencia

9) (x) (y) [x es un pedazo de uranio 235 (de tamaño mayor a uno determinado) e y es un pedazo de uranio 235 (del mismo tamaño) y x es impulsado contra $y \supset$ el Sol se pone púrpura brillante].

En esta discusión, yo mismo no sólo he mencionado sino que también he utlizado modalidades y quizá también condicionales fuertes. Debe considerarse que éstos sirven para enunciados metalingüísticos adecuados (en realidad meta-metalingüísticos), que sería difícil y más bien poco elegante establecer enteramente. El lector, si está interesado en ello, puede tratar de producir enunciados adecuados en el metalenguaje de mi presente discurso.

La presente sugerencia es, pues, que los condicionales subjuntivos y fuertes pueden ser sustituidos por afirmaciones relativas a lo que se sigue a partir de lo que está en una teoría bien comprobada. Estos enunciados metalingüísticos pueden ser expresados todos en un lenguaje extensional, utilizando el condicional débil "$\supset$". Ahora bien, puesto que el concepto de "teoría bien comprobada" no es formal, y su aplicación ha de dejarse al tacto y criterio de los científicos, nuestro idioma metalingüístico, que realiza la función de estos condicionales, no es, en cierto sentido, completamente formalizable. Por otra parte, debe observarse que una vez que decidamos qué teoría bien comprobada tenemos en cuenta, y si esta teoría está formalizada, entonces nuestras afirmaciones metalingüísticas pueden ser formalizadas, ya que la noción de "se sigue de" puede ser definida rigurosamente utilizando los métodos de la metamatemática (sintaxis).

Se dice, a veces, que aunque los condicionales fuertes no se necesiten en las teorías científicas, siguen siendo necesarios en el laboratorio. Supóngase que un científico se dice a sí mismo: "Si conecto los cables de mi aparato de este modo los quemaré, y si los conecto de esta otra manera recibiré una violenta sacudida eléctrica." Habiéndose dicho esto a sí mismo, decide no conectar los cables de *ninguna* de estas maneras, y por tanto, lo que dice, si se interpreta extensionalmente, es vacuamente verdadero. Por supuesto, no es la verdad vacua de esta sentencia lo que a él le interesa. Una vez más, sin embargo, podemos afirmar que lo que él

se está preguntando es si determinadas condiciones se siguen de su bagaje de conocimientos teóricos, y que está utilizando esta información metalingüística como ayuda para tomar decisiones prácticas. De este modo, si el científico sabe que a partir de una proposición *A* se sigue la proposición "recibiré una violenta sacudida eléctrica" hace todo lo que puede para asegurar la falsedad de *A*. En el laboratorio incluso, por consiguiente, parece ser que un científico puede arreglárselas con el condicional débil "⊃", junto con ciertos conceptos metalingüísticos expresables dentro de una lógica extensional.

Consideraciones semejantes ayudan a arrojar luz acerca del tema de las propiedades disposicionales. Se ha señalado frecuentemente que si, por ejemplo, "la sal (cloruro sódico) es soluble en el agua" se analiza como

10) (x) (x es cloruro sódico y x está inmerso en agua $\supset$ x se disuelve)

entonces es vacuamente verdadera si ninguna parte de cloruro sódico ha sido inmersa jamás, ni nunca lo será, en agua. Desde luego, el mar está lleno de cloruro sódico, y, evidentemente, el cloruro sódico, de hecho, ha sido inmerso en agua, pero algunos filósofos todavía desean tener lugar para afirmar que incluso en un universo en el que el cloruro sódico nunca ha sido y nunca será inmerso en agua podemos seguir discutiendo la cuestión, como algo no trivial, de si el cloruro sódico es soluble en agua o no. Podemos explicar parte de lo que estos filósofos desean indicando que 10) se sigue de las leyes de la naturaleza junto con determinadas sentencias que describen la estructura del cloruro sódico. Por supuesto 10) es un teorema de la teoría cuántica de la unión química. Además, incluso antes de que se hubiese pensado en la moderna teoría de valencias 10) podía ser construida como anticipo de alguna teoría de la solubilidad todavía sin desarrollar.

¿No necesitaría la teoría cuántica de la unión química postular ella misma propiedades disposicionales de cosas más fundamentales, tales como electrones? Sin embargo, algunos filósofos podrían desear ser capaces de afirmar que en un universo en el que no hubiese protones en absoluto, seguiría siendo cierto que *si hubiese* tantos protones como electrones estos *se atraerían* entre sí. Hay dos modos de explicar esta posición. Podría decirse que

11) (x) (y) (x es un protón e y es un electrón $\supset$ x e y se atraen mutuamente)

es un axioma de la teoría, y su posición como axioma es suficiente para proporcionarle un status no trivial, completamente aparte de cualquier cuestión relativa a la existencia de protones. En segundo lugar, podría señalarse que simplemente *no* desearíamos afirmar lo que serían los electrones si no existiesen los protones. Exactamente en este nivel fundamental la física y la cosmología están con toda seguridad demasiado íntimamente

relacionadas para que hagamos suposición alguna acerca de lo que serían las leyes de la naturaleza si los hechos del caso (condiciones iniciales o limitadoras) fuesen diferentes. Por ejemplo, si la teoría de los electrones y protones de J. A. Wheeler, como extremos opuestos de "agujeros de gusano" pertenecientes a un espacio-tiempo multiplemente-conectado hubiese de ser aceptada [21], ¿qué sentido tendría preguntar si los electrones tendrían carga de no existir protones? ¿Qué sentido tendría imaginar siquiera electrones sin protones?

En cualquier caso, si sustituimos enunciados relativos a disposiciones por enunciados extensionales en términos de cuantificadores universales y "$\supset$", esto nos llevará a falsedad cuando hablemos del mundo real. Las verdades triviales, tales como las de la forma "$(x)\,Fx \supset Gx$", cuando "$\sim(\exists x)\,Fx$" es verdadera, no pueden causar daño alguno si nos mantenemos dentro de una perspectiva extensional y no tratamos de discutir mundos no reales pero posibles. (Podemos dejar para ciertos metafísicos la preocupación acerca de si tiene algún sentido tal charla relativa a mundos no reales.) Si en realidad deseamos hablar ostensiblemente acerca de mundos posibles, podemos decir lo que necesitemos decir si utilizamos un idioma meta-teórico (metalingüístico), por ejemplo discutiendo la deducibilidad de determinadas sentencias de una teoría a partir de otras. Si no podemos decir lo que deseamos decir por medio de un idioma meta-teórico, entonces puede pensarse que, en realidad, no sabemos de lo que estamos hablando.

Sugerencias respecto a lecturas posteriores

Sobre medición fundamental y derivada, y conceptos relacionados con éstos, además de N. R. Campbell, *An Account of the Principles of Measurement and Calculation* (London: Longmans, Green, 1928) y *Foundations of Science* (New York: Dover, 1957), II parte, véase Brian Ellis, *Basic Concepts of Measurement* (New York: Cambridge University Press, 1966); C. W. Churchman y P. Ratoosh (eds.), *Measurement: Definitions and Theories* (New York: Wiley, 1959); y S. S. Stevens, "On the Theory of Scales of Measurement", *Science*, CIII (1946), 677-680. Sobre la inexactitud de las leyes físicas véase Michael Scriven, "The Key Property of Physical Laws— Inaccuracy" en Herbert Feigl y Grover Maxwell (eds.), *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart and Winston, 1961), págs. 91-101, con la réplica de H. Mehlberg, *ibíd.*, págs. 102-103, y la respuesta de Scriven, *ibíd.*, págs. 103-104. Véase también D. H. Mellor, "Experimental Error and Deducibility", *Philosophy of Science*, XXXII (1965), 105-122, que apoya también la crítica de Feyerabend y Sellars relativa a la relación jerárquica ortodoxa de la explicación científica, que fue discutida en el capítulo 3, anteriormente.

En la cuestión relativa al realismo *versus* instrumentalismo en la física, una discusión amplia desde un punto de vista más bien metafísico que neutral, puede verse en C. G. Hempel, "The Theoretician's Dilemma" en H. Feigl, M. Scriven y G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, Vol. II (Min-

[21] Véase pág. 183.

neapolis: University of Minnesota Press, 1958), págs. 37-98, y también en Ernest Nagel, *The Structure of Science* (New York: Harcourt, Brace and World, 1961), capítulo 6. Una posición que no es fácilmente clasificable como realista de forma no ambigua o instrumentalista de forma no ambigua, es la de Stephen Toulmin en su *Philosophy of Science* (London: Hutchinson, 153, y en rústica, New York: Harper and Row, 1960). Esta se apoya en parte en el concepto de Ryle de leyes de la naturaleza como "billetes de inferencia" o "licencias de inferencia". Véase G. Ryle, *Concept of Mind* (London: Hutchinson, 1949, y en rústica, New York: Barnes and Noble, 1950). Esta concepción de las leyes de la naturaleza ha sido criticada por H. Gavin Alexander, "General Statements as Rules of Inference?", *Minnesota Studies in the Philosophy of Science,* Vol. II, *op. cit.,* págs. 309-329. Para el operacionismo véase P. W. Bridgman, *The Logic of Modern Physics* (New York: Macmillan, 1927, y en rústica, 1960). Una defensa filosófica de la interpretación de Copenhague de la mecánica cuántica es dada por N. R. Hanson, "The Copenhagen Interpretation of Quantum Theory", *American Journal of Physics,* XXVII (1959), 1-15, y en forma resumida ha sido reeditada en A. Danto y S. Morgenbesser (eds.), *Philosophy of Science* (New York: Meridian Books, 1960). Véase también el trabajo de Hanson, "Five Cautions for the Copenhaguen Interpretation's Critics", *Philosophy of Science,* XXVI (1959), 325-337. Los intentos de K. R. Popper de una interpretación realista de la teoría cuántica pueden verse en el capítulo 9 de su *The Logic of Scientific Discovery* (London: Routledge and Kegan Paul, 1963, y en rústica, ed. rev., New York: Harper and Row, 1965, traducción española: *La lógica de la investigación científica,* ed. Tecnos, Madrid, 1962), y en su trabajo "The Propensity Interpretation of Probability and the Quantum Theory", y la discusión por parte de S. Körner (ed.) que le siguió, *Observation and Interpretation in the Philosophy of Physics* (New York: Dover, 1957), págs. 65-70 y 78-89. Sobre la teoría de la mecánica cuántica de la medición, en lo relativo a las cuestiones de realismo e instrumentalismo en la física, véase H. Margenau, "Measurements and Quantum States", *Philosophy of Science,* XXX (1963), 1-16 y 138-157, e Hilary Putman, "A Philosopher Looks at Quantum Mechanics", en R. G. Colodny (ed.), *Beyond The Edge of Certainty* (Englewood Cliffs, N. J.: Prentice-Hall, 1965), págs. 75-101. Sobre Berkeley como precursor de Mach en cuanto fenomenalista científico, véase el capítulo 6 de la obra de Popper, *Conjectures and Refutations* (New York: Basic Books, 1963).

Gran número de los trabajos de Feyerabend tratan de materias relevantes para la cuestión del realismo y el instrumentalismo en la física, pero en particular debe mencionarse su "Problems of Microphysics" en R. G Colodny (ed.), *Frontiers of Science and Philosophy* (Pittsburgh: University of Pittsburgh Press, 1962), páginas 189-283 y "Realism and Instrumentalism" en Mario Bunge (ed.), *The Critical Approach to Science and Philosophy* (New York: Free Press, 1964), págs. 280-308. Véase también J. J. C. Smart, *Philosophy and Scientific Realism* (London: Routledge and Kagan Paul, 1963), capítulo 2; Wilfrid Sellars, *Science, Perception and Reality* (London: Routledge and Kegan Paul, 1963, traducción española: *Ciencia, percepción y realidad,* ed. Tecnos, Madrid, 1971), capítulos 3 y 4; Grover Maxwell, "The Ontological Status of Theoretical Entities", en Herbert Feigl y Grover Maxwell (eds.), *Minnesota Studies in the Philosophy of Science,* Vol. III (Minneapolis: University of Minnesota Press, 1962), págs. 3-27; Marshall Spector, "Models and Theories", *British Journal for the Philosophy of Science,* XVI (1965), 121-142; Marshall Spector, "Theory and Observation", *British Journal for the Philosophy of Science,* XVII (1966), 1-20 y 89-104; R. Harré, *Theories and Things* (New York. Sheed and Ward, 1961); Max Born, "Physical Reality", *Philosophical Quarterly,* III (1953), 139-150; C. F. Presley, "Laws and Theories in the Physical Sciences", *Australasian Journal of Philosophy,* XXXII (1954), 79-103 [reeditado con las referencias omitidas en A. Danto y S. Morgenbesser (eds.), *Philosophy of Science* (New York: Meridian Books, 1960), págs. 205-255]; J. B. Thornton, "Scientific Entities", *Australasian Journal of Philosophy,* XXXI (1953), 1-21 y 73-100; y Joseph Agassi, "Sensationalism", *Mind,* LXXV (1966), 1-24. Acerca de la

cuestión histórica de si Einstein fue operacionalista o positivista, véase Robert Neidorf, "Is Einstein a Positivist?", *Philosophy of Science,* XXX (1963), 173-188.

Una exposición del teorema de Craig, con una discusión de su relevancia filosófica, se encuentra en el trabajo de Hilary Putnam "Craig's Theorem", *Journal of Philosophy,* LXII (1965), 251-260. Una réplica de E. Nagel a este trabajo aparece en las págs. 429-432 del mismo volumen de la revista, en la cual Nagel debate la discusión de Putnam del tratamiento del teorema de Craig en el libro de Nagel *The Structure of Science* (New York: Harcourt, Brace and World, 1961), páginas 134-137. Otra discusión del teorema de Craig puede verse en J. J. C. Smart, *Philosophy and Scientific Realism, op. cit.,* págs. 29-32. Sobre las leyes de la naturaleza en las lógicas extensionales y la cuestión de las propiedades disposicionales, véase H. Reichenbach, *Nomological Statements and Admissible Operations* (Amsterdam: North-Holland, 1954, y New York: Humanities Press, 1954); Wilfrid Sellars, "Counterfactuals, Dispositions and the Causal Modalities" en H. Feigl, M. Scriven y G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science,* Vol. II (Minneapolis: University of Minnesota Press, 1958), págs. 225-308, y Arthur Pap, "Disposition, Concepts and Extensional Logic", *ibíd.,* págs. 196-224; G. Bergmann, "Dispositional Properties", *Philosophical Studies,* VI (1955), 77-80; y "Comments on Professor Hempel's 'The Concept of Cognitive Significance'", *Proceedings of the American Academy of Arts and Sciences,* LXXX (1951), 78-86; W. V. Quine, *Word and Object* (Cambridge: Massachusetts Institute of Technology Press, 1960), especialmente págs. 222-225; y J. J. C. Smart, "Dispositional Properties", *Analysis,* XXII (1961), 44-46.

6. INDUCCION E HIPOTESIS

El problema filosófico de la inducción

En el capítulo precedente señalábamos a veces que se ha visto que son falsas determinadas teorías y surgía espontáneamente la idea de si podemos tener razón alguna para suponer que incluso las teorías mejor corroboradas no puedan correr algún día una suerte semejante. Es decir, supóngase que comprobamos una teoría T en n momentos distintos (y posiblemente incluso de n modos distintos) y que nunca resulta falsificada por nuestras observaciones, ¿constituye esto motivo alguno, o incluso una probabilidad no cero, de que no habrá pruebas posteriores que falsifiquen la teoría? ¿Tenemos alguna razón para suponer siquiera que la *próxima* prueba no falsificará la teoría? Es evidente que normalmente consideramos a las teorías bien comprobadas como probablemente verdaderas, pero la justificación teórica de esta actitud presenta problemas extraordinariamente difíciles de resolver. El problema de confirmar una teoría es un caso especial de la cuestión general (el problema tradicional) de la inducción: si hemos observado n As y hemos visto que todos ellos son Bs, y nunca hemos observado un A que no sea B, ¿nos da esto motivos para suponer, o siquiera una probabilidad no cero, de que todos los As son Bs, o siquiera que el próximo A será B? (En el caso especial considérese la propiedad de ser A como la propiedad de ser una prueba de una teoría particular, y la propiedad de ser B como la propiedad de ser una prueba confirmadora o no refutadora de la teoría.) Estas consideraciones sugieren que el problema tradicional de la inducción no puede resolverse simplemente señalando el hecho de que la ciencia no consiste en generalizaciones a partir de observaciones sino que hace uso del método hipotético-deductivo. El problema surge en otro nivel, cuando a partir del éxito pasado de una teoría generalizamos su probable verdad. De este modo, aunque los términos en los que se ha discutido normalmente el problema filosófico de la inducción (a saber, mediante referencia a generalizaciones del tipo de la historia natural, tales como "todos los cuervos son negros") no están en armonía con la imagen de la ciencia que se ha presentado hasta el momento en este libro, no será tan irrelevante como pudiera parecer a primera vista si comenzamos discutiendo la cuestión según la forma en que el problema ha surgido a partir de los escritos de David Hume [1].

El problema de la inducción se ha discutido tradicionalmente según el siguiente planteamiento: Supóngase que tenemos un número determinado de leyes de la forma "todos los As son Bs". Tales leyes se afirman porque, o bien un gran número de As examinados han sido Bs, y ninguno de los

*A*s examinados ha dejado de ser *B*, o porque las leyes han sido deducidas directa o indirectamente de leyes que han sido establecidas directamente del modo acabado de mencionar. Por ejemplo, sería razonable suponer que todas las muestras de determinada sustancia química tengan un espectro exactamente igual al observado en el caso de una muestra en particular, pero se sostiene que aquí nuestra confianza en la extrapolación a partir de un solo caso se basa en la inducción de que todas las muestras de la misma sustancia química muestran el mismo tipo de espectro, lo cual se basará en múltiples inducciones acerca de lo que se ve siempre que ocurre con multitud de sustancias diferentes. Finalmente, se dirá, volvemos a las inducciones directas, a las proposiciones de la forma "todos los *A*s son *B*s", y estas inducciones se basan en múltiples observaciones de *A*s en las que siempre hemos encontrado que eran *B*s. A esto debe añadirse la aclaración de que pueden hacerse inducciones en leyes que son de tipo estadístico, y en cuyo caso nuestras inducciones son de proposiciones de la forma "*m* de cada *n* *A*s observados han sido *B*s" a proposiciones de la forma "la probabiildad de que *A* sea *B* es *m*/*n*". Esta probabilidad se interpreta normalmente como una frecuencia límite. En el capítulo 2 he mantenido la interpretación de la tendencia de la probabilidad de Popper. Esta supera determinadas objeciones físicas y matemáticas a la noción de una probabilidad como frecuencia límite, pero no es de gran importancia a los efectos de la presente discusión, ya que tanto las proposiciones acerca de tendencias como las discusiones acerca de frecuencias límite se comprueban exactamente del mismo modo, a saber, mediante observaciones de las frecuencias observadas en un número bastante elevado. (A menos, por supuesto, que constituyan axiomas de una teoría, en cuyo caso pueden ser comprobadas indirectamente comprobando la teoría. Esta consideración no se plantea en el presente contexto.)

Hume demostró, de modo concluyente, que nuestra creencia en que todos los *A*s pasados, presentes y futuros son *B*s no puede depender de un argumento deductivo válido a partir del hecho de que todos los *A*s observados hasta el momento sean *B*s. No podemos deducir siquiera la proposición mucho más débil de que el *próximo A* observado será *B*. Su razonamiento es muy sencillo: No existe *contradicción* alguna en afirmar la proposición "todos los *A*s observados son *B*s, pero algunos de los *A*s todavía no observados no son *B*s". De modo semejante no puede haber ningún argumento deductivo válido a partir de las frecuencias observadas ("*m*/*n* de los *A*s observados han sido *B*s") a ninguna proposición fáctica relativa a la frecuencia de *A*s entre los futuros *B*s. No hay inconsistencia alguna en afirmar que *m*/*n* de los *A*s observados han sido *B*s y que ninguno, todos, o cualquier proporción intermedia que se quiera mencionar, de futuros *A*s serán *B*s.

Tampoco, como ha observado Hume, puede apoyarse la validez de la inducción en un argumento inductivo. Supóngase que hayamos observado que la mayoría de los argumentos inductivos del pasado han resultado válidos. Argumentamos entonces que la mayoría de los argumentos induc-

tivos, ahora y en el futuro, serán válidos. Este último argumento, que es un argumento inductivo acerca de los argumentos inductivos, puede ser denominado argumento inductivo de "segundo orden". Supóngase que su conclusión se establece de la siguiente forma: el principio inductivo es válido en argumentos de primer orden. El argumento hace uso de un principio inductivo de segundo orden, que es un *principio* o *regla* del argumento y no una *premisa* de él. Por tanto, como Max Black [2] y R. B. Braithwaite [3] han señalado, no hay *petitio principii* formal. Con todo, quien pudiera dudar de que los usos de primer orden del principio de inducción estuviesen justificados también tendría que dudar de la justificación del uso de segundo orden del principio. Si los usos pasados del principio fueron casos de argumento de a partir de "todos los *A*s observados han sido *B*s" a "todos los *A*s son *B*s", ¿cómo podríamos saber jamás que los usos pasados del principio inductivo estuvieron justificados o no, ya que *ex hypothesi* no sabemos si los *A*s no examinados son *B*s? Debemos considerar, por consiguiente, que nuestro argumento de segundo orden se refiere a argumentos de primer orden de la forma: "Todos los *A*s examinados hasta el momento han sido *B*s; por consiguiente, el próximo *A* examinado será *B*." Podemos comprobar, en realidad, que en determinados casos la inferencia ha quedado justificada por el éxito: el *próximo A* examinado resultó ser *B*. Podríamos entonces proponer el argumento de segundo orden de que puesto que (quizá) la mayoría de los argumentos de la forma "todos los *A*s son *B*s; por consiguiente, el próximo *A* será *B*" han tenido éxito, así también el próximo uso del argumento inductivo de primer orden tendrá él mismo, también, éxito. Este argumento no sería circular en el sentido de que su conclusión fuese una de sus premisas, pero no conseguiría convencer a quien se sintiese escéptico con relación al valor del principio inductivo, ya que se sentiría tan escéptico con relación a este segundo uso como con relación al primero.

Podríamos añadir que aunque se aceptase un argumento inductivo de segundo orden acerca de la validez de los argumentos inductivos en la vida ordinaria, no parece que existan muchas posibilidades de aportar una defensa inductiva de las inducciones *científicas,* en el sentido de los argumentos para la validez de teorías bien comprobadas. Pues se da el caso de que casi todas las teorías bien comprobadas del pasado han demostrado ser, posteriormente, incorrectas en sentido estricto. Lo más que podría lograr un argumento inductivo de segundo orden sería la validez de las teorías en cuanto instrumentos de predicción en determinadas circunstancias. (Así, aunque la teoría de la gravitación de Newton es falsa, sigue siendo un instrumento de predicción del todo útil para los fabricantes de almanaques náuticos.) El apoyo inductivo de la ciencia parecería ser, por tanto, una política más atractiva para el instrumentalista científico que para el realista, ya que aunque la mayor parte de las teorías científicas no han resultado ser estrictamente verdaderas, casi todas han continuado siendo útiles mecanismos predictivos.

Otra forma en que los filósofos han intentado contestar a la argumen-

tación de Hume ha sido proponiendo algún presupuesto fáctico relativo al universo que, de ser verdadero, podría justificar la inducción. El más famoso de ellos es el Principio de la Variedad Limitada de J. M. Keynes [4]. Para ver cómo operaría éste sustituyamos, en primer lugar, la inducción por la simple enumeración a través de la inducción eliminativa. Consideramos a la propiedad de ser *A* como un miembro de un conjunto finito de propiedades P_1, ..., P_m, *A*, cada una de las cuales *podría* ser condición suficiente para que algo fuese *B*. Se supone también que no existe ninguna otra propiedad, fuera del conjunto, que pudiese ser condición suficiente para que algo sea *B*. Entonces, encontrando un P_1, que no sea *B*, luego un P_2 que no sea *B*, y así sucesivamente, agotamos todas las posibilidades excepto *A*, lo cual nos permite fijar el ser *A* como condición suficiente para ser *B*, es decir, el afirmar "todos los *A*s son *B*s". Esta técnica puede extenderse evidentemente para tratar acerca de las condiciones necesarias también, ya que "*A* es una condición suficiente para *B*" es equivalente a "*B* es una condición necesaria para *A*".

Es evidente que esta técnica de inducción eliminativa depende de presupuestos fácticos relativos al universo. Hace falta que sepamos que el conjunto finito P_1, P_2, ..., P_m, *A*, agota todas las propiedades relevantes que hay que considerar. Podría argumentarse que en tanto en cuanto el conjunto P_1, P_2, ..., P_m, *A* no contenga una proporción *inperceptiblemente pequeña* de la totalidad de propiedades relevantes, el procedimiento de eliminar P_1, P_2, ..., P_m proporciona alguna probabilidad no cero de que *A* es una condición suficiente de *B*. Este es aproximadamente el método de J. M. Keynes. El postula un Principio de Variedad Limitada que asegura que el conjunto de propiedades que han de ser eliminadas es finito. Pero, una vez más, el dilema de Hume alza su fea cabeza. El Principio de la Variedad Limitada es, evidentemente, un presupuesto fáctico con relación al mundo, y por tanto, *a priori*, no puede saberse si es verdadero, mediante puro razonamiento. Por otra parte, se necesita para justificar la inducción, y, por tanto, no puede ser él mismo justificado por el argumento inductivo.

La misma dificultad surge con relación a otros posibles principios de razonamiento inductivo. Así pues, supóngase que se dice que es una "necesidad del pensamiento" el que la misma causa produzca el mismo efecto. En realidad *no* se trata de una necesidad del pensamiento, ya que es rechazada en la mecánica cuántica. Esta consideración muestra que la alegada necesidad del pensamiento es un principio empírico, y que ha sido refutado (si creemos en el establecimiento de la mecánica cuántica). A fin de establecerlo tendría que ser establecido inductivamente, y en este caso no podría ser un principio en el que descanse todo el razonamiento inductivo. Podría ser considerado como un principio heurístico: un requerimiento para continuar buscando causas cuando todavía no se ha descubierto causa alguna. De ser así, resulta ser fácticamente vacuo. Desde luego, si la mecánica cuántica es correcta sería un requerimiento que en algunos casos pudiera ser que nunca condujese al éxito.

En respuesta al dilema de Hume algunos filósofos modernos han adop-

tado la posición de que la inducción no *necesita* ser justificada. Por ejemplo, Paul Edwards, en un famoso escrito, "Bertrand Russell's Doubts about Induction" [5], argumenta que en la ciencia "buenas razones" *significan* "buenas razones inductivas"; es decir, que, en última instancia, la ciencia se apoya en el principio de que si todos y sólo los *A*s se ha observado que son *B*s, es legítimo considerar esto como *alguna* razón para afirmar "todos los *A*s ahora y en el futuro son *B*s". Supóngase que un filósofo plantea la cuestión de si los científicos tienen buenas razones para pensar que el futuro será igual que el pasado en determinados aspectos relevantes, por ejemplo, que si una persona salta sin paracaídas u otra ayuda de un piso alto del edificio "Empire State" esa persona, al igual que cualquier otra que haya saltado de un edificio alto en el pasado, se caerá con rapidez en el suelo. ¿Qué mejor razón, se pregunta Edwards, puede darse que el que este tipo de cosas ha ocurrido en el pasado? Podemos señalar la teoría de la gravitación, pero esto, simplemente, nos lleva de nuevo a una extrapolación de a partir de predicciones pasadas de la teoría de la gravitación que han tenido éxito a suponer que la teoría será válida en el futuro. El tipo de razones que los científicos *sí* dan para suponer (en algún caso concreto) que el futuro se conformará a las pasadas experiencias *son* precisamente lo que designamos comúnmente por "buenas razones inductivas".

Edwards ilustra esta cuestión mediante una analogía [6]. Ordinariamente la palabra "physician" se usa en inglés de un modo un tanto vago: puede considerarse que significa, por ejemplo, "una persona que se ha graduado en medicina", o puede considerarse, quizás, más ampliamente como "persona que posee la habilidad de curar enfermedades en un grado considerablemente superior al del hombre medio". Sin embargo, aunque se usa de un modo un tanto vago, *se usa* dentro de límites bien definidos: por ejemplo, sería incorrecto denominar a un experto reparador de bicicletas "physician". Edwards imagina ahora que alguien hace la paradójica observación de que en realidad no hay "physicians" en New York. Aunque él cree que está contraviniendo la opinión general sobre el asunto, resulta que con "physician" él no alude a nada semejante a "persona que se ha graduado en medicina y que posee una habilidad para curar enfermedades superior a la normal" sino que alude a una "persona que se ha graduado en medicina y que puede curar cualquier enfermedad imaginable en menos de dos minutos". En principio podríamos considerar que tal persona está haciendo una interpretación fáctica que está en desacuerdo con la nuestra propia, pero si resultase que tal persona se negase a llamar "physician" incluso a quien hubiese de curar todas las enfermedades imaginables en unos meses, o en cualquier período de dos minutos o más de duración, entonces veríamos que nuestro interlocutor estaba simplemente utilizando la palabra "physician" de un modo no habitual. Indica también Edwards que Russell, al decir que no tenemos ningún motivo para suponer que el futuro se parecerá al pasado, no está haciendo ninguna afirmación real. Es evidente, por lo que él dice, que nada que ca-

rezca de un argumento deductivo válido que parta de las premisas y llegue a una conclusión les satisfará, y, como Hume ha demostrado, ninguna deducción válida tal es posible. Russell ha modificado el uso de la expresión "buenas razones" de modo que lo que el científico normal denomina "buenas razones" ya no cuenta como tal concepto, exactamente igual que la persona que utiliza "physician" del modo peculiar que acabamos de explicar anteriormente, no permite que ninguna de las personas a las que *nosotros* denominaríamos de buen grado "physicians" cuenten como "physicians". Ni esta persona ni Russell han hecho ninguna afirmación sorprendente, sino que han efectuado un cambio en su uso de las palabras.

Tal "disolución" del problema de la inducción parece atractivo a primera vista, e incluso a segunda vista podemos admitir que ayuda a enfocar el problema y a disipar ciertos confusionismos que nos podrían hacer pensar en el escéptico filosófico con relación a la inducción, que no se satisface con *ninguna* consideración empírica, como en una especie de científico ultra-precavido, a quien casi ninguna consideración empírica puede satisfacer. El científico ultra-precavido ni siquiera se aproxima a la posición del escéptico filosófico, al igual que un número elevado, por ejemplo 10^{100}, no se aproxima siquiera al número transfinito $\aleph_0$. Hasta aquí todos de acuerdo. Obsérvese, sin embargo, que podrían defenderse exactamente igual métodos científicos para predecir el futuro, o para intentar llegar a la verdad con relación al mundo, mediante una disolución lingüística del problema de su validez. De este modo, si afirmamos que buenas razones científicas son lo que los científicos normalmente denominan buenas razones, así también, de igual modo, buenas razones mágicas son lo que los magos normalmente denominan buenas razones.

Aunque estemos de acuerdo con Edwards, y filósofos post-wittgensteinianos de mentalidad semejante, en que el escéptico con relación a la inducción que desee una justificación deductiva está pidiendo la lógicamente imposible, no tenemos por qué admitir que una vez que hemos visto esto tengamos que concluir que se trata de un pseudo-problema debido a confusión lingüística. Quizás exista otro planteamiento del problema de justificar la inducción. Podemos admitir que la inducción no puede ser "validada" (validated); que es lógicamente imposible demostrar que la conclusión de una inducción se sigue lógicamente de la evidencia empírica. Podríamos abandonar este intento, al igual que los matemáticos han desertado en el intento de cuadrar el círculo, debido a una prueba algebraica de la imposibilidad de llevarlo a cabo. Con todo, quizá podamos "vindicar" la inducción, usando la terminología de Herbert Feigl. Feigl [7], siguiendo a Hans Reichenbach [8], señala la posibilidad de una forma de argumento cuya conclusión no es la proposición "el método inductivo de predecir el futuro sirve", sino, más bien, la proposición más débil "si hay algún método de predecir el futuro que sirva, entonces la inducción sirve". Obsérvese que el argumento pretendería mostrar no que si sirve cualquier método entonces la inducción es el *mejor* método, sino, sim-

plemente que *si* sirve cualquier método (y nosotros no sabemos si sirve o no) *entonces* la inducción es *uno* de estos métodos. (O de modo equivalente, por contraposición, mostrar que si la inducción no sirve, entonces ningún otro método servirá tampoco.)

El dilema de Hume no muestra que el intento de vindicar la inducción esté condenado al fracaso, ya que un argumento vindicativo no tiene por objeto demostrar que la inducción es una política válida. Su objeto es, simplemente, demostrar lo razonable de la inducción en el sentido de ser un método a adoptar tan bueno como el mejor. El argumento se parece más bien a la apuesta de Pascal, que tiene por objeto demostrar no que el Cristianismo es verdadero, sino que es razonable actuar de acuerdo con el presupuesto de que es verdadero [1].

Si el argumento de Edwards demuestra que alguien que trate de *validar* la inducción está en la posición del posible cuadrador del círculo, esto por sí mismo (en un sentido) no resuelve el problema de justificar la inducción a menos que a su vez mostremos la imposibilidad matemática de una *vindicación* de la inducción, también. Si pudiéramos demostrar que tanto la validación como la vindicación de la inducción son imposibles, podría alegarse que esto resolvería el problema de la inducción, en un sentido semejante que la prueba algebraica acerca de la imposibilidad de la cuadratura del círculo, mediante regla y compás únicamente, sirvió para cancelar este problema. Jerrold J. Katz ha seguido, últimamente, esta línea de pensamiento [12]. Pero incluso si se pudiera prescindir del mismo modo del problema de la posibilidad de vindicación de la inducción, uno todavía podría preguntarse si no podría aparecer algo que pudiera ser considerado como una justificación pero que no fuese ni una prueba de validez ni una prueba de vindicación. Podría tratarse de un tipo de justificación en la que nadie haya pensado todavía. Después de todo, fue, comparativamente, hace muy poco cuando se postuló la idea de la vindicación. ¿Cómo podemos estar seguros que no sufra un nuevo cambio toda la cuestión de la justificación de la inducción, un día de estos? Se trata, por lo menos, de una posibilidad que tener en cuenta, puesto que la cuestión no es de las que puedan ser formuladas con precisión, al igual que ocurría en el caso de la cuadratura del círculo.

A fin de determinar la posibilidad de vindicar la inducción será conveniente concretar la discusión echando un vistazo a alguno de los argumentos de Wesley Salmon, que ha sido uno de los más ardientes defensores de la idea de vindicar la inducción [13]. Este autor ha formulado la cuestión del modo siguiente: Considérese una secuencia de acontecimientos de un tipo determinado, y supongamos que una proporción f_n de los n primeros acontecimientos posee una determinada propiedad. Nuestro problema consiste en descubrir cuál es la proporción a la que f_n tiende,

[1] Por supuesto, no pretendo defender el argumento de Pascal, que es falaz. Véase Antony Flew, *Is Pascal's Wager the Only Safe Bet?*" [9] y James Cargile, "*Pascal's Wager*" [10]. Las diferencias entre la apuesta de Pascal y el intento de vindicar la inducción pueden verse en un escrito de Isabel Creed [11].

conforme n tiende a infinito. Si utilizamos la política deductiva habitual conjeturaremos que $\lim_{n\to\infty} f_n = f_n$. Esta es la denominada regla "directa", y es una regla "convergente". Otras reglas "convergentes" son de la forma: Si f_n es la frecuencia observada para n acontecimientos observados, conjetúrese que $\lim_{n\to\infty} f_n = f_n + c_n$, siendo c_n alguna función especificada de n (distinta que la función $c_n = 0$) tal que $c_n \to 0$ conforme $n \to \infty$. Tales reglas se denominan reglas "oblicuas". Si presuponemos que el propósito de la inducción es obtener frecuencias límite (lo cual puede ser cuestionado, desde luego) entonces podemos demostrar que la política de adoptar cualquiera de las mencionadas reglas convergentes puede ser vindicada, al menos como política para seres inmortales, ya que podemos demostrar que, de existir una frecuencia límite, la adopción de tal regla convergente, ya sea directa u oblicua, nos llevará finalmente a adentrarnos en cualquier intervalo pre-asignado (por pequeño que sea) que circunde a esta frecuencia límite.

Es fácil comprobar la superioridad de cualquiera de estas reglas convergentes frente a algunas otras reglas. Considérese, por ejemplo, la política "contra-inductiva" de suponer que $f_n \to \frac{n-m}{n}$ (siendo m, igual que antes, el número de muestras positivas de observaciones n). La gente, *en realidad,* a menudo piensa bastante de acuerdo con esta regla. Si yo, jugando con otra persona, arrojo un penique al aire y gano las cuatro primeras tiradas, mi oponente puede tener la esperanza de ganar en la quinta ocasión. (En Australia algunos redactores deportivos al parecer piensan de este modo cuando predicen los resultados de arrojar la moneda para elegir los primeros turnos del "batsman" en los partidos de cricket, lo cual es de gran importancia en este juego.) Es cierto que la gente que hace tales predicciones contra-inductivas no sigue en realidad la regla contra-inductiva, sino que argumenta a partir de una confusa idea con relación al curso de la buena suerte. En realidad, por supuesto, si arrojamos un penique bien estable, entonces mi probabilidad de ganar en la quinta tirada es de un medio, con independencia de lo que haya ocurrido en las anteriores tiradas, pero, desde luego, si no sabemos si los peniques son bien estables o no, y los resultados de las anteriores tiradas constituyen toda nuestra evidencia, la evidencia de cuatro "caras" sucesivas sugiere, de acuerdo con la política inductiva habitual, que el penique no está estable y que lo más probable es que no tengamos "cara" en la quinta ocasión también. La superioridad de la política inductiva frente a la contra-inductiva puede verse como sigue: Si la secuencia (f_n) tiene un límite l, entonces el seguir la política inductiva nos permitirá, finalmente, acercarnos cada vez más a él, mientras que la política contra-inductiva nos acercará arbitrariamente a un valor distinto, a saber, $1-l$. Otra dificultad relativa a la política contra-inductiva es que puede llevar a contradicción, del siguiente modo: Supóngase que lanzamos repetidamente un dado bien estable. La regla inductiva ordinaria nos llevaría, desde luego,

al valor $\frac{1}{6}$ como la frecuencia límite de una cara determinada (por ejemplo la cara "dos"). La regla contra-inductiva nos llevaría a postular $\frac{5}{6}$ como la frecuencia límite de aquella misma cara. Ahora bien, puesto que hay seis caras, y las frecuencias límites postuladas para cada una de ellas serían las mismas, a saber $\frac{5}{6}$, tendríamos que afirmar que la frecuencia de aparición de cualquiera de las seis caras sería $6\times\frac{5}{6}=5$. Esto es un absurdo, ya que no se puede dar significado a ninguna frecuencia mayor que uno. En cualquier caso, es una cuestión de pura lógica el que la frecuencia de tiradas en las que aparezca una u otra cara sea igual a uno. (Podemos considerar que las tiradas caprichosas en las que el dado cae en una ranura del suelo, por ejemplo, y se tumba hacia un lado o una esquina no cuentan como tiradas legítimas.)

Considérese ahora la siguiente regla convergente no inductiva, que no lleva a contradicciones de la forma que fue descrita en el parágrafo anterior. Supóngase que distinguimos seis posibles resultados de la operación de tirar un dado y que como resultado conjeturamos una frecuencia límite de $\frac{1}{6}$ para cada uno de los posibles resultados (un "uno", un "dos", un "tres", y así sucesivamente). A ésta Salmon la denomina regla "*a priori*". Es evidente que esta regla es convergente $\left(\text{para } \frac{1}{6}\right)$ pero que no converge para la frecuencia límite real, excepto en el caso de un dado bien confeccionado. Una regla interesante que puede demostrarse que proporciona convergencia para la frecuencia límite, si existe, en todos los casos, es la de asignar una frecuencia $\frac{1}{n+1}\cdot\left(m+\frac{1}{k}\right)$, siendo k el número de alternativas lógicamente posibles (por ejemplo 6 en el caso de las tiradas de un dado). Obsérvese que para $n=0$, es decir, antes de toda experiencia, asigna el valor $\frac{1}{k}$, como ocurría con la regla *a priori*. Esta regla es un término medio entre la regla *a priori* y la regla inductiva directa. Por supuesto es una de las reglas oblicuas, ya que es de la forma "asígnese f_n+c_n como frecuencia límite", siendo $c_n\to 0$, cuando $n\to\infty$. La expresión $\frac{1}{n+1}\cdot\left(m+\frac{1}{k}\right)$ puede escribirse de la forma $\frac{m}{n}+\frac{n-mk}{kn\,(n+1)}$ y $\frac{n-mk}{kn\,(n+1)}\to 0$, cuando $n\to\infty$.

Presuponiendo que todas las observaciones precedentes son correctas, el problema que le queda por resolver a Salmon es el de justificar el uso de la regla que asigna f_n (la regla directa) en vez de la infinidad de reglas oblicuas que asignan f_n+c_n, siendo $c_n\to 0$, cuando $n\to\infty$. Resulta que

todas las asignaciones de las funciones c_n de las reglas oblicuas tienen la propiedad de ser lenguaje dependiente. Por ejemplo, la regla intermedia del parágrafo precedente depende del número de posibilidades alternativas k. Pero supóngase que tuviésemos un lenguaje que distinguiese únicamente tres caras del dado, de tal modo que "un 'uno' o un 'dos'", "un 'tres' o un 'cuatro'", "un 'cinco' o un 'seis'" correspondiesen a los predicados últimos de este lenguaje. Quienes hablasen este lenguaje no distinguirían seis sino tres posibles resultados de la tirada de un dado, y tendríamos $k=3$, no 6 como antes. Salmon utiliza el criterio de la permanencia lingüística para desechar todas las funciones c_n, excepto aquélla para la cual c_n es siempre igual a cero. Sin embargo, Stephen Barker [14] ha objetado que incluso la regla de la inducción (para la cual $c_n=0$) viola el requisito de la permanencia lingüística. Trae a colación el caso de los predicados "grue" y "bleen" de Goodman [15] (véase pág. 53 anteriormente). Se recordará que definíamos "grue" con el significado de "verde con anterioridad al año 2000 y azul con posterioridad". Por tanto, hasta el momento todas las esmeraldas han sido a un tiempo verdes y "bleen", y por tanto la regla de Salmon justifica la inducción de "todas las esmeraldas son "bleen"", no menos que la de "todas las esmeraldas son verdes" [2]. Como hemos observado ya (pág. 53) de nada servirá negar que los predicados "grue" y "bleen" no son predicados científicos propios, ya que sus definiciones mencionan la fecha específica, año 2000. Puede alegarse que "grue" y "bleen" podrían ser considerados como los predicados no definidos, en cuyo caso "blue" (azul) definido como "'bleen' hasta el año 2000 y 'grue' posteriormente" y "green" (verde) definido como "'grue' hasta el año 2000 y 'bleen' posteriormente", serían aquéllos cuya definición implicaría una específica referencia al tiempo. Salmon sugiere, por tanto, una nueva restricción, que es la de que los predicados básicos de un lenguaje científico deben ser predicados "puramente ostensivos", es decir, predicados que puedan ser definidos ostensiblemente y tales que las semejanzas entre sus muestras sean susceptibles de inspección directa.

Esta réplica a la paradoja de Goodman no servirá para aquellos que acepten la posición general que se adopta en este libro, según la cual no existe ningún lenguaje de observación básico, y también según la cual las sentencias observacionales no se pueden distinguir de las no observacionales simplemente mirando su vocabulario. Desde luego, desde el punto de vista adoptado en este libro, el acudir a una práctica científica inductiva mediante la vindicación del método de la inducción por medio de simple enumeración parece artificial desde un principio, al menos, quizá, cuando

---

[2] Hablando estrictamente, la regla de Salmon no puede justificar la inferencia de una conclusión universal, ya que unos cuantos *A*s podrían ser *B*s y la frecuencia de los *B*s entre los *A*s seguir tendiendo a 1 conforme el número de *A*s tiende al infinito. Sin embargo, esto carece de importancia, ya que podemos o bien dejar de afirmar "Todos los *A*s son *B*s" y decir, simplemente, que la frecuencia límite de los *B*s entre los *A*s es 1, o de lo contrario modificar la regla especificando que una conclusión universal puede derivarse únicamente si no se ha observado que *ningún A* fuese *B*.

se aplica a la metateoría de la ciencia, es decir, a argumentar, pasando de los pasados aciertos predictivos de teorías bien comprobadas a los futuros aciertos predictivos de teorías bien comprobadas. (Es difícil imaginar cómo podría ser utilizado para argumentar a favor de la *verdad,* en oposición a los aciertos predictivos generales de las teorías, ya que casi todas las teorías del pasado han resultado ser, en un sentido estricto, falsas.) Sin embargo, el intento de Salmon no deja de ofrecer interés, ya que proporciona una muestra de un intento cuidadosa y seriamente calculado para vindicar la inducción; muestra lo que uno generalmente piensa cuando habla de un intento de *vindicar* la inducción, a la vista de la prueba de Hume de la imposibilidad de *validarla.* Desde luego, Salmon ha mostrado un modo de progresar respecto a la posición de Humen, que al propio Hume le hubiera resultado atractivo, de haberlo conocido, ya que el propio Hume creía, en efecto, que la ciencia se basa, en última instancia, en la inducción por medio de simple enumeración y creía, también, que existía un conjunto de predicados básicos que tan sólo podían ser definidos ostensiblemente [3].

Habiendo observado algunos de los detalles de un importante intento de vindicar la inducción, volvamos a considerar la idea de la vindicación de un modo más general. Considérese cualquier política *P* para conseguir un tipo determinado de resultados *R.* Quizá no podamos demostrar que *P* es válida, es decir, que consigue R, pero quizá podamos demostrar que de no conseguirse *R* con *P* entonces no se conseguirá con ninguna otra política tampoco. En cuyo caso, confiando en que será la mejor medida, posiblemente consideremos sensato adoptar *P.* Aplicando esto a la ciencia, Salmon ha considerado que *R* era la predicción de las frecuencias límite, acerca de secuencias infinitas de acontecimientos. Mas, puesto que no somos inmortales, no podemos aplicar el método inductivo más que durante un determinado intervalo de tiempo limitado. Demostrar que si cualquier método es válido *en última instancia* entonces la inducción lo es, no significa, en modo alguno, demostrar que si cualquier método es válido dentro de un cierto intervalo de tiempo fijado (por ejemplo 100 años) entonces la inducción lo es [18]. Es cierto que Salmon ha considerado el problema de "The Short Run" ("A corto plazo") en un trabajo que lleva ese mismo título [19], pero en él se ocupa únicamente del problema de predecir frecuencias a corto plazo, mientras que nuestro problema actual es distinto: se trata de predecir frecuencias a largo plazo cuando sólo tenemos frecuencias a corto plazo de las que partir.

Quizá salgamos mejor parados si intentamos vindicar alguna cosa distinta de la inducción mediante simple enumeración. Después de todo, si

---

[3] Este apoyarse en la ostensión sugiere que Salmon cree en este contexto que es posible separar un conjunto especial de predicados observacionales. Incluso de ser así, él en realidad lo rechaza en un trabajo más reciente [16], que no he podido ver hasta que ya era demasiado tarde para discutirlo en el presente volumen. En aspectos más importantes tal escrito muestra que Salmon ha modificado su posición, en parte en respuesta a las críticas de Ian Hacking [17].

observamos la práctica científica real no vemos que los científicos se ocupen de conjeturar frecuencias límite basándose en frecuencias observadas a corto plazo. Salmon es del todo consciente, desde luego, de esto; él justifica su propia posición como una útil "idealización".

Observemos la práctica inductiva científica real. Como ha indicado Brian Ellis en un artículo muy interesante [20], aun cuando hayamos tirado un dado con anterioridad, si arrojásemos un dado homogéneo y simétrico 100 y obtuviésemos 55 "caras" no esperaríamos que la frecuencia se aproximase a 0,55 después de muchas más pruebas. Dejaríamos a la suerte la diferencia en 0,05 y predeciríamos una frecuencia de 0,50. Obrar de otro modo supondría ir en contra de teorías bien establecidas de la geometría y la mecánica. En la práctica, por consiguiente, nuestras predicciones están "afectadas por la teoría" y nunca nos encontramos aplicando la regla directa, excepto, probablemente, a la luz de consideraciones teóricas con las que podemos ver que no está en conflicto. Incluso una generalización de la historia natural tal como "todos los cuervos son negros" no es el resultado de la simple inducción a partir de "todos los cuervos observados son negros", sino que depende de presupuestos a la base, relativos a la herencia de las características dentro de las especies animales y a las consecuentes semejanzas en el color de las aves dentro de una sola especie. Es en este punto donde aparecen las observaciones de Salmon relativas a que su sistema está basado en una "idealización". La ciencia desarrollada, podría argüirse, depende de una concatenación de aplicaciones de la regla directa. Ellis replica que si la ciencia fuese cuestión de aplicación de la regla directa, entonces la formulación de hipótesis científicas sería una cuestión puramente mecánica y podría ser realizada por computadores, cuando, por el contrario, la experiencia sugiere que requiere métodos no efectivos[4], dependiendo de la imaginación y el ingenio. No estoy seguro de que en este punto su argumento sea formalmente convincente, ya que un método que depende de una concatenación de procedimientos efectivos no tiene por qué ser efectivo si no existe ninguna regla efectiva para decidir en qué orden han de ser aplicados estos procedimientos efectivos. Sin embargo, aun en este caso, la formulación de hipótesis científicas no se asemeja, ciertamente, en lo más mínimo, a una concatenación de aplicaciones de la regla directa de la inducción, aunque sólo sea debido a que implica la invención de nuevos conceptos.

La posición de Ellis respecto a que la dependencia teórica es una condición necesaria para la inducción científica le permite proporcionar una respuesta simple a la paradoja de Goodman. La mayoría de nosotros rechazaría el uso de los predicados "grue" y "bleen" debido a que nuestras teorías no nos llevan a suponer que exista nada especial en el primero de enero del año 2000; y, además, si todas las esmeraldas *se cam-*

[4] En el sentido técnico de la lógica matemática. En el sentido ordinario, ¡desde luego que la ciencia es completamente efectiva! En el sentido técnico, decir que los procedimientos científicos son efectivos sería afirmar que pueden ser aplicados automáticamente, al igual que un procedimiento de decisión en lógica.

biasen, efectivamente, de verde a azul, y viceversa, en tal fecha, esto iría en contra de nuestras teorías actuales de óptica y química física. Esta idea de la dependencia universal de la teoría está en armonía con la posición adoptada en este libro y con la idea de K. R. Popper de que antes de que se invente una hipótesis debe existir una hipótesis previa, ya que la observación presupone siempre alguna hipótesis, a la luz de la cual se realiza. Popper [21] ha comparado la cuestión de "¿qué es lo que aparece primero, la hipótesis o la observación?" con la cuestión "¿qué es lo que aparece primero, el huevo o la gallina?". La respuesta a esta última es: "un huevo previo", y a la primera es: "un tipo de hipótesis previo". Popper ha argumentado que incluso gran parte de la psicología animal se comprende mejor como una cuestión de formación de hipótesis rudimentarias más que como una cuestión de aprendizaje mediante reflejos condicionados (que correspondería a la inducción por medio de simple enumeración).

En esencia, y sin hacer del todo justicia a las complejidades con que es elaborado, el argumento de Ellis parece reducirse a esto: la predicción científica requiere una dependencia teórica, y por tanto no existe otro camino mejor que el de usar teorías ya existentes, o, en caso de que pudiera demostrarse tras posteriores observaciones que estas teorías llevaban a predicciones incorrectas, utilizar, entonces, nuevas teorías que pueda verse que son por lo menos tan acertadas como lo han sido las anteriores. Ahora bien, a mi parecer Ellis ha dado una buena razón para la predicción dependiente de la teoría en lugar de recurrir a la regla directa de Salmon. Con todo, también parece que no ha hecho nada para justificar el método científico frente a un escéptico radical; uno que desease saber por qué los métodos científicos deben *seguir* siendo acertados, y por qué, a partir de la media noche en adelante, por ejemplo, no sería mejor la brujería o el arte de la adivinación. Es cierto que si la brujería o el arte de la adivinación comenzasen a ser predictivamente superiores a la ciencia, aparecerían grandes dificultades, basándonos en las teorías científicas existentes, para explicar por qué las cosas de pronto comenzaban a comportarse de un modo tan diferente. Pero esto no es de esperar, con toda seguridad, que tuviese peso suficiente para alguien que estuviese de antemano dispuesto a considerar la posibilidad de que los métodos absolutamente no científicos tendrían mayor eficacia predictiva.

Otra nueva consideración es la siguiente: hemos observado que la vindicación hace siempre relación a un fin o a un resultado propuesto. Ellis, al igual que Salmon, presupone que esto es *predicción,* aunque continúa hacia el final de su trabajo tratando con la cuestión fundamental de la aceptación de las teorías. Señala que sería irracional aceptar o rechazar cualquier teoría únicamente basándonos en lo que *podría* ocurrir. (Mas el escéptico con relación a la inducción cuestionará esto; él *rechazará,* o, al menos, *se negará a aceptar* las teorías, a causa de su preocupación de que pudieran darse ejemplos contrarios.) Entonces, si la experiencia nos ha hecho rechazar una teoría T_1 (por ejemplo la teoría calórica), mien-

tras que otra teoría T_2 (por ejemplo la teoría cinética del calor) *puede* ser aceptada todavía, y todavía no ha sido propuesta ninguna otra teoría que pueda ser aceptada, Ellis mantiene que debe ser racional aceptar T_2. Esto sugiere una política semejante a la que ha sido propuesta por Hilary Putnam [22].

La explicación de Putnam implica una excesiva simplificación de la práctica científica real, pero se acerca más a la ciencia real que el tipo de concepción de la ciencia sugerido en las discusiones de Reichenbach y Salmon. La idea de Putnam viene a ser la siguiente: consideramos únicamente hipótesis "efectivas". Ligeramente simplificado, el método de Putnam es el siguiente: Supóngase que *M* es una propiedad de determinada clase de individuos. Entonces "una hipótesis efectiva sobre M" es aquélla a partir de la cual podemos calcular de un modo mecánico si un individuo posee la propiedad *M* o no. Denomínese a las hipótesis sobre *M* que sobrevivan en un momento *m*, "las alternativas". (Si durante las pruebas de comprobación de las alternativas en cualquier momento particular m_0, se sugieren nuevas hipótesis sobre *M*, éstas se añaden al conjunto de las alternativas.) Calculamos si el próximo individuo observado tiene o no tiene la propiedad *M*. Conservamos las alternativas que predicen el resultado (*M* o no *M*, según el caso) y rechazamos aquéllas que predicen el resultado contrario. Continuamos hasta que todas las alternativas, excepto una, han fallado, o hasta que todas lo han hecho. Si queda sólo una exactamente la aceptamos como verdadera.

Obsérvese que puede ser que estemos equivocados al aceptar la hipótesis que queda como verdadera. Nuevas pruebas pueden llevarnos a rechazarla también. Con todo, el método presenta la siguiente deseable propiedad: si formulamos una hipótesis que es a la vez efectiva y verdadera, entonces el método nos llevará a aceptar la hipótesis si es verdadera [5]. La idea de Putnam parece constituir una vindicación parcial de la inducción, ya que es, con toda probabilidad, imposible que el defensor de los métodos absolutamente no científicos (que, con toda probabilidad, no implican las pruebas observacionales de las hipótesis) demuestre que *sus* métodos presentan una propiedad comparable. Con todo, no constituye una vindicación completa de la inducción, ya que no nos permite demostrar que si un método cualquiera es capaz de permitirnos encontrar hipótesis verdaderas, este método nos lo permitirá. No tenemos seguridad alguna de que lleguemos a pensar alguna vez en la hipótesis verdadera acerca de *M*, y tampoco tenemos seguridad alguna de que podremos es-

[5] En el caso de hipótesis estadísticas acerca de *M*, una hipótesis abandonada puede ser re-establecida. Si obtenemos 550 "caras" a partir de 1.000 tiradas de un penique, podríamos abandonar la hipótesis de que la frecuencia es ½ como "demasiado improbable". Pero 5.001 "caras" en las siguientes 10.000 tiradas nos llevará a rechazar nuestro rechazo precedente. Con todo, considero que el método de Putnam será válido para hipótesis probabilísticas, así como también para las no probabilísticas, en el sentido de que *si* alguna vez se encontrase la hipótesis verdadera ésta permanecería. (Sobre las reglas condicionales del rechazo de hipótesis de frecuencia, véase R. B. Braithwaite, *Scientific Explanation*, tradución española: *La explicación científica*, ed. Tecnos, Madrid, 1964, capítulo 6 [23].)

tablecer la hipótesis verdadera en un tiempo razonable, aun en el supuesto condicional de que un método u otro nos permitirá hacer esto. Evidentemente, de nada servirá si la observación que hubiera de eliminar una alternativa ocurrirá sólo después de millones de años. Al señalar que el método de Putnam no nos muestra cómo vindicar del todo la inducción, no estoy postulando esto, sin embargo, como una crítica al escrito de Putnam, ya que su método aparece en el curso de una discusión en la que se ocupa de otras materias.

En cualquier caso, las hipótesis en la ciencia real no son siempre efectivas. (Recuérdese que el propio Putnam mantuvo que estaba únicamente sugiriendo una simplificación del método científico.) Aparte del hecho de que en general no existe seguridad alguna de que descubramos alguna vez una prueba de un teorema demostrable en una teoría física (aunque este tipo de no efectividad no es muy importante en la práctica) no existe seguridad ninguna de que la presencia de un caso aparente de no *M* pueda ser considerada como refutación ni siquiera de una hipótesis no probabilística sobre *M*. Esto es así debido a que, como Pierre Duhem ha señalado [24], en la ciencia puede salvarse una teoría a la vista de aparente refutación mediante alguna modificación en alguna hipótesis auxiliar. Es decir, cuando comprobamos una hipótesis *H* investigando alguna consecuencia observacional *O*, la situación es, normalmente, de que *O* no se sigue de *H* únicamente, sino de *H* junto con ciertos supuestos auxiliares *S*. En la astronomía planetaria, por ejemplo, *H* puede ser la mecánica newtoniana, mientras que *S* contendrá enunciados de "condiciones iniciales" (posiciones de los planetas previamente observadas) y también enunciados de óptica, ya que es mediante medios ópticos como se averiguan las posiciones planetarias. De este modo, si encontramos a un planeta en una posición en la que no debiera estar según nuestra teoría de la mecánica, es formalmente posible conservar las leyes de la mecánica y rechazar nuestros supuestos ópticos y explicar la discrepancia postulando algún efecto óptico que implique el descenso de rayos de luz desde el planeta. Desde luego, no se haría esto, pero se precisa de tacto y habilidad para saber cuándo rechazar y cuándo no las hipótesis auxiliares, y es este aspecto de la ciencia el que impide que las hipótesis reales sean "efectivas". Por otra parte, como Ellis ha señalado [25], el continuo abandono de las hipótesis auxiliares llevaría a intolerables complicaciones en la práctica, y con frecuencia no es del todo desencaminado el considerar las hipótesis científicas como si fueran efectivas en el sentido de Putnam.

El método de Putnam, acabado de describir, es perfectamente bien una formalización del método científico, tal como fue descrita por K. R. Popper. Popper, sin embargo, parece considerar cualquier intento de justificar la inducción como desencaminado [26]. De hecho considera la inducción como un mito. Si por "inducción" entiende la inferencia inductiva de los textos de lógica, quizá tenga razón, pero si por "inducción" se entiende la predicción del futuro basándose en el pasado, mediante teorías científicas,

no se trata, evidentemente, de un mito, si no de algo que hacemos siempre. Yo, por tanto, continuaré disponiendo para mi uso de los conceptos de "inducción" y "procedimiento inductivo" al discutir la posición de Popper. Popper, en efecto, mantiene que si cambiamos el fin de los inductivistas, que desean justificar los procedimientos inductivos, por otro, a saber, el *crítico* (criticism), es evidente que si las hipótesis falsas son refutables mediante experiencias adecuadas, entonces el método científico puede ser vindicado con sólo consideraciones de la lógica deductiva. De hecho, se podría decir que no puede ser validado nunca, ya que la falsedad de una hipótesis puede ser deducida de una observación adecuada simplemente por *modus tollens,* es decir, un argumento de la forma "$H \supset O$ y $\sim O$; por consiguiente $\sim H$".

Popper rechaza las teorías de la confirmación dadas por Carnap [27] y filósofos de ideología semejante. Esto no resulta sorprendente, ya que en tales teorías de la confirmación, la confirmación de una ley universal h mediante evidencia e es dada mediante $c(h, e)=0$. Por ejemplo, supóngase que h es "todos los As son Bs", y supóngase que se considera que $c(h, e)$ es el radio del número de las As ya observadas al número de todas los As posibles. Si consideramos que el número de todos los As posibles es infinito, entonces $c(h, e)=0$. Carnap tiene una definición más complicada y sutil de $c(h, e)$, pero sigue proporcionando el valor 0 cuando h es una hipótesis universal. Carnap supera esta dificultad introduciendo la noción de "confirmación de muestras". La confirmación de muestras de una hipótesis h es la confirmación de la hipótesis de que la *próxima* muestra observada del fenómeno en cuestión estará conforme con h. La teoría de la confirmación de Carnap puede proporcionar un valor no cero a la confirmación de muestras de una hipótesis universal, aun cuando el valor de la propia función de confirmación sea cero. Podemos afirmar, entonces, que la confirmación de una ley ha de ser considerada como su confirmación de muestras.

Sin embargo, este recurso a la confirmación de muestras parece presuponer una filosofía instrumentalista o positivista de la ciencia. Si una teoría es considerada simplemente como un útil instrumento de predicción, entonces la confirmación de las muestras será lo que tendrá valor para nosotros. Si nos interesa, principalmente, comprender el universo, entonces la cuestión será distinta, ya que es posible que una hipótesis falsa tenga una amplia confirmación de muestras. Considérese una ley tal como "todos los As son Bs" y supongamos que hasta que transcurran cien millones de años a partir de ahora no hay ningún A que no sea B. Supongamos que la confirmación de muestras es elevada, como en el caso presupuesto ocurrirá probablemente. Un pragmatista no necesitaría saber nada más, en realidad, ya que la elevada confirmación de instancias le llevaría a falsas espectativas sólo cuando transcurran un millón de años a partir de ahora, y no es probable que por entonces, en fecha tan lejana del futuro, él anduviese todavía por ahí. El realista, sin embargo, no estará satisfecho, ya que él deseará saber si *todos* los As son Bs o no,

incluyendo incluso las que estén distantes de nosotros en el espacio y el tiempo. Debe observarse también, como una grave objeción a la teoría de la confirmación de Carnap, que sólo puede ser aplicada a lenguajes bastante simples, que no son lo bastante ricos para expresar la ciencia real.

Ahora bien, Popper también desea mantener que la confirmación [6] de *h,* dado *e,* es cero cuando *h* es una hipótesis universal. Sin embargo, él desea mantener que algunas hipótesis son mejor *corroboradas* que otras, no entendiendo por "corroboración", "confirmación", en el sentido "de hacer más probable". (Tampoco, dicho sea de paso, desea Carnap identificar su "confirmación" con "hacer más probable", ya que si *h* es una ley universal, ninguna suma de evidencia *e* hará que *c*(*h, e*) sea otra cosa que cero. Para Carnap es la confirmación de *muestras* lo que puede incrementarse mediante nueva evidencia). Popper desea mantener que una hipótesis puede ser mejor corroborada que otra aun cuando sus probabilidades sean cero en ambos casos. Es decir, una de las hipótesis puede ser sometida a *pruebas de comprobación más severas* que la otra. La severidad de una prueba de comprobación es un concepto que no puede ser formalizado: no es cuestión simplemente del número de pruebas de una hipótesis. Implica la noción de *sinceridad* por parte del experimentador u observador y de sus colegas en las cuestiones teóricas que le piden que realice determinadas observaciones. Deben de tratar de modo sincero y activo de comprobar la teoría en aquellas circunstancias que les parezca menos probable que resulte válida, así como en aquellos casos en los que se espera que sea acertada. Deben tratar de encontrar refutaciones de la hipótesis, no solamente "comprobaciones" o "confirmaciones" [7] de ella. La corroboración se diferencia de la confirmación en que cuanto menos probable es una hipótesis mejor puede ser corroborada. En otras palabras, superar una prueba difícil es más valioso que superar una fácil. Popper ha sugerido que ciertas hipótesis psicoanalíticas pueden ser corroboradas en muy escasa medida (si es que pueden serlo en alguna) ya que es difícil, o tal vez imposible, concebir experimentos u observaciones que, de darse de cierto modo, pudiesen ser consideradas por los psicoanalistas como refutadores de su teoría [28]. Las teorías físicas son generalmente corroboradas en alto grado; levantan su cuello muy alto, por así decirlo, y a veces le es posible al físico especificar algunos resultados anteriormente sorprendentes que pueden ser comprobados experimentalmente. Popper cita el ejemplo de la teoría general de la relatividad de Einstein, que predecía la bajada de un rayo de luz de una estrella lejana, conforme el rayo pasaba cerca del Sol. Estas predicciones fueron comprobadas durante el eclipse de 1917 en las expediciones a Sobral, Brasil, por A. C. D. Crommelin y C. Davidson, y a Príncipe, en el Golfo de Guinea, por A. S. Ed-

---

[6] La confirmación es una noción metalingüística. Guarda íntima relación con el concepto metalingüístico de probabilidad. (Véase el capítulo 2, pág. 39.)

[7] Escribo estas palabras entre comillas, ya que, por supuesto, de acuerdo con Popper no se puede dar realmente ni una comprobación ni una confirmación de una hipótesis científica.

dington y E. T. Cottingham. (La relatividad especial también predice una bajada, pero sólo la mitad de la de la relatividad general.) La teoría de Einstein corrió un *riesgo* al predecir este efecto, ya que de no haberse dado tal efecto hubiera tenido que abandonar Einstein la teoría general de la relatividad. Tal vez Popper no haya elegido, en este caso, el mejor ejemplo, en vista de las considerables dificultades para valorar las pruebas de la relatividad general (tanto las mencionadas aquí como otras, tales como la del adelanto del perihelio de Mercurio, que fue predicho por la teoría). De hecho es la dificultad de comprobar la teoría general de la relatividad (hasta la fecha por lo menos) lo que ha hecho que se le considere como una teoría más bien especulativa o metafísica, a diferencia de la teoría especial de la relatividad de Einstein, de la que hay pruebas abundantes, tanto ópticas como dinámicas (tales como las observaciones de colisiones entre partículas aceleradas y sus trayectorias).

Podría pregutarse cómo es posible que Popper diga, como dice, que la hipótesis *más improbable* es corroborada por una prueba válida, si todas las hipótesis universales tienen la misma probabilidad (cero). La respuesta es que Popper tiene una teoría de la *estructura refinada* de la probabilidad [29], de modo que si una hipótesis h_2 es deducible de otra hipótesis h_1, pero h_1 no es deducible de h_2, podemos decir que h_1 es más probable que h_2, aun cuando $c(h_1, e)=c(h_2, e)=0$.

Queda por resolver, sin embargo, la cuestión de por qué nos debe interesar el comprobar (y tal vez corroborar) las teorías, si no podemos justificar la atribución de una probabilidad mayor de cero ni siquiera a la teoría mejor corroborada. Una de las respuestas que Popper ha dado es la siguiente: si una teoría h_1 está mejor corroborada que h_2, debido a que supera todas las pruebas a las que h_2 ha sido sometida y también algunas pruebas que h_2 no ha superado, entonces se trata de una teoría más *fuerte que* h_2. Por tanto, el fortalecimiento y las pruebas continuas de nuestras teorías conducen siempre a verdades más observacionales. Desde luego, aunque concibamos una nueva prueba que falsifique nuestras teorías actuales, el resultado de este experimento aumenta nuestros conocimientos de lo que ocurre en las diversas circunstancias [30]. Esto supone, ciertamente, una buena justificación para buscar corroboraciones, pero sólo desde un punto de vista instrumentalista, que es, en lo fundamental, tan inaceptable para Popper como para mí. De no estar muy interesados en las consecuencias observacionales, salvo en tanto en cuanto son pruebas de teorías, ¿no necesitamos, entonces, alguna justificación de la corroboración en términos de la verdad, o al menos de la probabilidad, de las teorías? Se diría, por consiguiente, que Popper no ha disipado la aparente necesidad de la vindicación de la inducción en las líneas generales de los intentos discutidos con anterioridad en este capítulo.

Objeciones de Duhem a la asimetría entre confirmación y refutación

La filosofía de la ciencia de Popper descansa, en gran medida, en una asimetría entre la comprobación y la refutación. Ninguna consecuencia no observada de una teoría puede comprobar la teoría, ya que esto supondría cometer la falacia lógica de argumentar pasando de $A \supset B$ y B a A. Por otra parte, la refutación de una teoría es posible, ya que el paso de $A \supset B$ y $\sim B$ a $\sim A$ es válido. Observábamos en la pág. 149, anteriormente, sin embargo, que Duhem se ha opuesto, en efecto, a esta asimetría, ya que mantenía que a la vista de la evidencia aparentemente refutadora, se podía salvar una teoría desechando algunos presupuestos auxiliares. Si H es una teoría y O una "consecuencia observacional", O no se sigue de H únicamente, sino de $H \cdot A$, siendo A ciertos presupuestos extraños a la propia H. Entonces, pues, por lo que a la lógica se refiere, a la vista de O tenemos la posibilidad de rechazar H o rechazar A. Así, cuando se observó que el planeta Urano se desviaba de la órbita calculada en la mecánica newtoniana, teniendo en cuenta las perturbaciones debidas a los planetas entonces conocidos, esto se consideró como una refutación de la mecánica newtoniana. Los supuestos auxiliares A (incluyendo datos acerca de los cuerpos perturbadores del sistema solar) fueron modificados por J. C. Adams y U. J. J. Leverrier (independientemente) para incluir el supuesto de la existencia de un planeta todavía desconocido con una órbita exterior a Urano. Luego predijeron la posición de este planeta, que posteriormente se comprobó que existía.

La respuesta de Popper a esto es que todos nuestros supuestos pueden ser cuestionados, pero que no podemos cuestionarlos todos al mismo tiempo. Podemos comprobar H, suponiendo que A no presenta problemas, pero en otro momento podemos comprobar A. Esto implicará supuestos auxiliares B, que pueden ser, a su vez, cuestionados con la ayuda de supuestos auxiliares C, los cuales pueden ser, a su vez, cuestionados, y así sucesivamente. De hecho, para Popper estas consideraciones, en cierto modo, confirman su posición falible. Dentro de la ciencia todo es arriesgado, incluso la aceptación de presupuestos de base (y esto también incluye la aceptación de lo que retiene nuestra memoria, y de nuestras percepciones sensibles).

Adolf Grünbaum ha señalado [31] que el esquema $(H \cdot A) \supset O$, $\sim O$, $\therefore H \vee \sim A$ es de interés solamente si es posible pensar en alguna hipótesis no trivial A' que implique $\sim A$. De lo contrario, no nos será posible modificar nuestras hipótesis auxiliares de un modo adecuado. No es evidente que exista una hipótesis no trivial tal. (Una A' *trivial* puede encontrarse siempre, por supuesto. Así podría servir $H \supset \sim O$, ya que $[(H \cdot H \supset \sim O) \supset \sim O]$ es verdadera.) Considérese, por ejemplo, el caso del adelanto del perihelio de Mercurio, que puede ser explicado mediante la mecánica newtoniana más los presupuestos de base habituales. En algún

momento se propuso salvar la mecánica newtoniana mediante la postulación de un planeta (al que se le llamó incluso, transitoriamente, "Vulcano") entre Mercurio y el Sol. Sin embargo, se demostró que tal hipótesis no sería válida. En este caso, por consiguiente, no es evidente que sea posible una modificación no trivial de los presupuestos auxiliares, que pudieran salvar a la mecánica newtoniana. (Estoy simplificando con exceso el caso, y descuidando ciertas dificultades que se dan al evaluar la significación de los datos observacionales. En conjunto, sin embargo, quedará patente que *probablemente* los datos no puedan ser explicados mediante presupuestos que estén de acuerdo con la mecánica newtoniana.) Grünbaum argumenta incluso que puede demostrar un caso en el que no es ni siquiera *lógicamente* posible que exista un tal A'. Si la noción de "lógicamente posible" significaba simplemente "que no contradiga las leyes de la lógica clásica" resultaría una noción suficientemente clara, pero, de hecho, Grünbaum requiere para su argumentación una distinción entre hechos y convenciones en geometría, lo cual sería desechado, probablemente, por la mayoría de los defensores de la posición de Duhen, por ejemplo W. V. Quine, contra el cual argumenta Grünbaum específicamente.

Conclusión

Parecerá evidente que el método científico, incluso considerado como medio de predicción, requiere, al efecto, de una vindicación, cuando no de una validación. Desde luego, si se aceptasen pruebas, tales como las sugeridas por Katz, de la imposibilidad matemática de esto, tendríamos que abandonar la búsqueda. Con todo, ¿no sentiríamos cierta insatisfacción, que *no* es sentida por el hombre que ve la imposibilidad matemática de la cuadratura del círculo? Podemos vivir perfectamente bien sin la cuadratura del círculo, pero en la práctica no podemos arreglárnoslas sin la ciencia como medio de predicción, y algunos de nosotros no nos sentimos felices sin la ciencia como medio de conseguir la verdad con relación al mundo. El simple rechazo de las falsedades no nos parece tan satisfactorio a alguno de nosotros como se lo parece a Popper, y nos gustaría algún método acerca del modo en el que sucesivas teorías científicas, aun cuando a su debido tiempo tengan que ser rechazadas, no sólo resultan más acertadas como medios de predicción, sino que, en cierto modo, nos proporcionan un "retrato más correcto de la realidad". Parece ser que distamos mucho de esto.

Con cierto alivio, pues, abandonemos ahora estas dudas escépticas y estas preocupaciones metafísicas, y aceptando, justificadamente o no, la verdad en general de las teorías del momento presente, veamos si la ciencia puede ayudarnos en la discusión de algunos de los problemas que han ocupado a los filósofos.

Sugerencias respecto a lecturas posteriores

Buenas introducciones filosóficas al problema de la inducción son G. H. von Wright, *The Logical Problem of Induction,* segunda edición revisada (Oxford: Blackwell, 1957, y New York: Barnes and Noble, 1965); William Kneale, *Probability and Induction* (Oxford: Oxford University Press, 1949); y Stephen Barker, *Induction and Hypothesis: a Study of the Logic of Confirmation* (Ithaca, N. Y.: Cornell University Press, 1957). En su *Treatise on Probability and Induction* (London: Routledge and Kegan Paul, 1951), Von Wright se ocupa más del análisis de la lógica formal que en su otra obra mencionada anteriormente. Una visión de los puntos de vista vigentes pueden obtenerse en el simposium editado por Henry E. Kyburg, Jr., y Ernest Nagel, *Induction: Some Current Issues* (Middletown, Conn.: Wesleyan University Press, 1963). Una utilísima revisión del trabajo en lógica inductiva desde 1951 es proporcionada por Henry E. Kyburg, Jr., "Recent Work in Inductive Logic", *American Philosophical Quarterly,* I (1964), 249-287. Incluye una bibliografía detallada que es un útil complemento de la bibliografía de la obra de Von Wright *The Logical Problem of Induction, op. cit.,* que incluye una buena lista de trabajo previo sobre la materia. Una útil bibliografía selecta sobre la inducción se encuentra en las págs. 169-171 de Paul Edwars y Arthur Pap (eds.), *A. Modern Introduction to Philosophy,* edición revisada (New York: Free Press, 1965). Críticas a algunos detalles de la obra de Kneale (*op. cit.*) son hechas por Jonathan Bennett, "Some Aspects of Probability and Induction", I y II, *British Journal for the Philosophy of Science,* VII (1956-57), 220-230 y 316-322.

Para la posición que mantiene que la inducción requiere una limitación en la variedad de la naturaleza, véase C. D. Broad, "On the Relation Between Induction and Probability", *Mind,* XXVII (1918), 389-404, y XXIX (1920), 11-45, y J. M. Keynes, *Treatise on Probability and Induction* (London: Macmillan, 1921, y en rústica, New York: Harper and Row, 1962). Otros postulados metafísicos para justificar la inducción son sugeridos por Bertrand Russell en su *Human Knowledge* (New York: Simon and Schuster, 1948, y en rústica, 1962).

Sobre la vindicación de la inducción, véase H. Reichenbach, "On the Justification of Induction", *Journal of Philosophy,* XXXVII (1940), 97-103, y *Experience and Prediction* (Chicago: University of Chicago Press, 1938), págs. 348-363; H. Feighl, "De Principiis Non Disputandum...?", en Max Black (ed.), *Philosophical Analysis* (Englewood Cliffs, N. J.: Prentice-Hall, 1950), págs. 119-156, y "On the Vindication of Induction", *Philosophy of Science,* XXVIII (1961), 212-216; muchos escritos de Wesley Salmon, especialmente "Inductive Inference" en Bernard Baumrin (ed.), *Philosophy of Science, The Delaware Seminar,* Vol. 2 (New York: Interscience, 1963), págs. 341-370, y "The Foundations of Scientific Inference" en Bernard Baumrin (ed.), *Philosophy of Science, The Delaware Seminar,* Vol. 2 (New York: Interscience, 1963), págs. 341-370, y "The Foundations of Scientific Inference" en R. G. Colodny (ed.), *Mind and Cosmos* (Pittsburgh: University of Pittsburgh Press, 1966), págs. 135-275. En este último escrito Salmon admite una crítica que hace Ian Hacking a sus intentos de vindicar la regla directa. Véase Ian Hacking, "Salmon's Vindication of Induction", *Journal of Philosophy,* LXII (1965), 260-266, que está escrito en gran medida desde el punto de vista de la estadística matemática. Una discusión del escrito de Hacking es la que hace I. Levi, "Hacking Salmon on Induction", *ibíd.,* 481-487. Una posición bastante distinta respecto a la vindicación de la inducción puede verse en Brian Ellis, "A Vindication of Scientific Inductive Practices", *American Philosophical Quarterly,* II (1965), 296-304.

J. J. Katz, *The Problem of Induction and Its Solution* (Chicago: University of Chicago Press, 1962), argumenta que el problema de la inducción puede ser disuelto demostrando que una justificación de la inducción, ya mediante valida-

ción o vindicación, es matemáticamente imposible. Argumentos relativos a que el denominado problema de la inducción es un pseudo-problema, que puede anularse mediante análisis "lingüístico" o "conceptual", pueden verse en unos cuantos capítulos de la obra de Max Black, *Language and Philosophy* (Ithaca, N. Y.: Cornell University Press, 1954), y en *Models and Metaphors* (Ithaca, N. Y. Cornell University Press, 1962, traducción española: *Modelos y metáforas*, ed. Tecnos, Madrid, 1967), Dentro de la misma línea filosófica están el trabajo de Paul Edwards "Bertrand Russell's Doubts About Induction" en A. Flew (ed.), *Logic and Language*, primera serie (Oxford: Blackwell, 1951, y en rústica, New York: Doubleday Anchor Books, 1965), págs. 55-79, y el capítulo 9 de la obra de P. F. Strawson, *Introduction to Logical Theory* (New York: Willey, 1952, y en rústica, New York: Barnes and Noble, 1952). J. O. Urmson, "Some Questions concerning Validity" en A. Flew (ed.), *Essays in Conceptual Analysis* (London: Macmillan, 1956, y New York: St. Martin's Press, 1956), págs. 120-133, admite y critica a la vez la intentada disolución del problema mediante un análisis lingüístico del tipo mencionado anteriormente. Véase también F. L. Will "Will the Future be like the Past", en A. Flew (ed.), *Logic and Language*, segunda serie (Oxford: Blackwell, 1953, y en rústica, New York: Doubleday Anchor Books, 1965), págs. 32-50, con crítica a cargo de D. Williams, "Induction and the Future", *Mind*, LVII (1948), 226-229. El trabajo de David Hume "Sceptical Doubt concerning the Human Understanding", que es la sección IV de su *Inquiry Concerning Human Understanding* (publicado por vez primera en 1748), se discute desde el punto de vista de la moderna filosofía analítica en el capítulo 4 del libro de A. Flew, *Hume's Philosophy of Belief* (London: Routledge and Kegan Paul, 1961, y New York: Humanities Press, 1961).

Justificaciones *a priori* de la inducción son intentadas por Donald C. Williams, *The Ground of Induction* (Cambridge: Harvard University Press, 1947) y Roy Harrod, *The Foundations of Inductive Logic* (New York: Harcourt, Brace and World, 1956).

La filosofía de la inducción de Rudolf Carnap puede verse especialmente en su *Logical Foundations of Probability*, segunda edición (Chicago: University of Chicago Press, 1962), así como también las págs. 966-979 de su "Replies and Expositions", en P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (La Salle, Ill.: Open Court, 1963). Este último volumen también contiene valiosos ensayos a cargo de J. G. Kemeny, A. W. Burks, H. Putnam y E. Nagel acerca de diversos aspectos de la obra de Carnap relativa a la lógica inductiva, y las réplicas de Carnap. Véase también G. H. von Wright, "Carnap's Theory of Probability", *Philosophical Review*, LX (1951), 362-374. La posición de K. R. Popper puede verse en su *Logic of Scientific Discovery* (London: Routledge and Kegan Paul, 1959, y en rústica, ed. rev., New York: Harper and Row, 1965, traducción española: *Sociología de la investigación científica*, ed. Tecnos, Madrid, 1962), y en su *Conjectures and Refutations* (London: Routledge and Kegan Paul, 1963, y New York: Basic Books, 1963), especialmente el capítulo 1. Sobre la interpretación y la crítica de Popper hay un animado intercambio de trabajos entre D. C. Stove (criticando a Popper) y J. W. N. Watkins (defendiendo a Popper) en artículos aparecidos en el *Australasian Journal of Philosophy*, XXXVII (1959) y XXXVIII (1960). El último volumen también incluye una nota crítica de Stove relativa a la obra de Popper, *Logic of Scientific Discovery, op. cit.* Véase también J. A. Passmore, "Popper's Account of Scientific Method", *Philosophy*, XXXV (1960), 326-331. Un tratado escrito desde un punto de vista afín al de Popper, es la obra de J. O. Wisdom *Foundations of Inference in Natural Science* (London: Methuen, 1952).

Nelson Goodman se ocupa mucho del problema de la inducción en su *Fact, Fiction and Forecast* (Cambridge: Harvard University Press, 1955, y en rústica, New York: Bobbs-Merrill, 1965). Y en particular destaca su perplejidad acerca de "grue" y "bleen". Esto ha sido muy discutido en diversidad de obras. Una reciente discusión es el simposium, "The New Riddle of Induction", con trabajos de Richard C. Jeffrey, Judith Jarvis Thomson, John R. Wallace y prefacio

y comentario a cargo de Nelson Goodman, *Journal of Philosophy*, LXIII (1966), 281-331. Véase también R. J. Butler, "Messrs. Goodman, Green and Grue", en R. J. Butler (ed.), *Analytical Philosophy*, segunda serie (Oxford: Blackwell, 1965), páginas 181-193.

Sobre la posición de Pierre Duhem en relación con la falsificabilidad de las teorías científicas, véase su *Aim and Structure of Physical Theory* (Princeton: Princeton University Press, 1954, y en rústica, New York: Atheneum, 1962) y la obra de Armand Lowinger, *The Methodology of Pierre Duhem* (New York: Columbia University Press, 1941). El argumento de Duhem vuelve a aparecer en la obra de Quine, *From a Logical Point of View*, segunda edición revisada (Cambridge: Harvard University Press, 1961, y en rústica, New York: Harper and Row, 1963), capítulo 2. Pueden verse críticas en A. Grünbaum, *Philosophical Problems of Space and Time* (New York: Alfred A. Knopf, 1963), capítulo 4; y A. Grünbaum, "The Falsificability of a Component of a Theoretical System", en P. K. Feyerabend y G. Maxwell (eds.), *Mind, Matter and Method, Essays in Philosophy and Science, in Honor of Herbert Feigl* (Minneapolis: University of Minnesota Press, 1966), páginas 273-305. Algunas manifiestas paradojas que surgen en relación con la inducción son discutidas hábilmente por Carl G. Hempel, "Recent Problems of Induction", en R. G. Colodny (ed.), *Mind and Cosmos, op. cit.*, págs. 112-134.

7. ESPACIO Y TIEMPO

ALGUNAS TEORÍAS FILOSÓFICAS ACERCA DEL ESPACIO

¿De qué tipo de objeto hablamos cuando hablamos del espacio y el tiempo? ¿Son ellos, o la combinación de ellos, una entidad sustancial, o se trata de conjuntos de relaciones entre cosas distintas a ellos mismos? ¿O son, quizá, formas que la mente impone en un mundo no espacial y no temporal como Inmanuel Kant mantenía? Veamos en este capítulo qué resulta de estos viejos problemas filosóficos cuando se contemplan a la luz de las ideas científicas. Comenzaremos por discutir el espacio, en la medida que pueda hacerse con independencia del tiempo.

Algunos pensadores griegos se inclinaban a imaginar las cosas en el espacio igual que peces nadando en el mar. Así Platón, en el *Timeo,* hablaba del espacio como de un "receptáculo". Esto preludiaría la que más tarde habría de ser denominada teoría "absoluta" del espacio. *A grosso modo* lo que son los peces con relación al mar eso son las cosas con relación al espacio. El espacio es el elemento último en el que existe el agua, o el aire, o la materia pesada. El hueco entre los trozos de la materia es el espacio vacío, y por tanto no es una nada, al igual que no lo es el mar que está entre peces diferentes. Al igual que el mar que está entre los peces es agua, así el "hueco" entre trozos de materia es espacio. Por supuesto que el espacio es más penetrante que el agua del mar, ya que el agua no penetra en el pez, y, en cambio, en la teoría absoluta del espacio, no se trata únicamente de que los trozos de materia estén en el espacio, sino de que los pedazos del espacio están en los trozos de materia: el espacio penetra por completo todo lo que está *en* el espacio. La teoría absoluta del espacio aparece en una forma muy pura en la filosofía de René Descartes. Según Descartes la esencia de la materia es la extensión, y, en consecuencia, el problema consistía, para él, en distinguir entre la materia y el espacio vacío. Si la esencia de la materia es la extensión, entonces, puesto que es manifiesto que el espacio vacío es extenso, se sigue de ello que no existe un auténtico vacío y que todas las partes del espacio están llenas de materia. Esto da lugar a un nuevo problema: ¿Cómo podía distinguir Descartes entre la materia ordinaria y la materia más sutil del espacio aparentemente vacío? ¿Qué explicación, por otra parte, podía dar del movimiento de la materia *en* el espacio? Veremos posteriormente que las dificultades de Descartes a este respecto pueden ser, al menos parcialmente, superadas mediante ideas matemáticas que no habían sido desarrolladas en su época.

Contra la teoría absoluta del espacio mantenida por Descartes, Gottfried

Leibniz propuso una teoría relacional. Según Leibniz el mundo estaba constituido enteramente de entidades no espaciales, llamadas "mónadas", y el espacio se analiza en términos de las relaciones entre estas mónadas. A partir de Leibniz, la teoría relacional del espacio ha sido propuesta normalmente sin la teoría de las mónadas o los otros portadores no espaciales de las relaciones espaciales. Según esta versión, el espacio es analizado en términos de relaciones entre cosas que están ellas mismas *en el espacio*. Ahora bien, si estas cosas que están en el espacio son espaciales y tienen extensión y figura, entonces se presenta una dificultad para comprender la teoría relacional que no estaba presente en la versión de Libniz de la teoría. Es bastante fácil considerar, por ejemplo, una serie de relaciones de parentesco: padre, madre, hijo, hija, sobrino, sobrina, primo, prima, nieto, nieta, etc. Las partes de padre o nieto no tienen relación de parentesco entre sí. De acuerdo con la teoría relacional, las relaciones que dan lugar a espacio son relaciones entre cosas, cuyos trozos están ellos mismos relacionados mediante relaciones dentro del conjunto. No estoy seguro de que ésta sea una dificultad en la teoría relacional del espacio. Quizá se trate más bien de una dificultad cuando nuestra física es una física de continuos, mientras que con una teoría de partículas moderna, las partículas últimas no son espaciales en ningún sentido. Ni tampoco, sin embargo, parecen ser simplemente no espaciales, como las mónadas de Leibniz. En cualquier caso, la teoría relacional del espacio es, en este aspecto, al menos, más fácil de establecer si suponemos con Leibniz que en último lugar se dan relaciones espaciales entre cosas no espaciales [1].

A pesar de esta duda metafísica acerca de si la teoría relacional del espacio necesita entidades no espaciales como *relata* últimas, la teoría relacional es en muchos aspectos muy plausible. Así, la distancia entre dos estrellas se considera fácilmente como una relación entre las estrellas y alguna medida de longitud. Nos encontraremos con otros argumentos a favor de la teoría relacional del espacio y el tiempo, pero, a fin de tratarlos contextualmente, será conveniente echar un vistazo primero a la posición de Isaac Newton con relación al espacio y el tiempo.

La teoría del espacio y el tiempo de Newton

En la "Glosa a las Definiciones", de los *Principia* [2], Newton distinguía el "tiempo absoluto, verdadero y matemático" del "tiempo relativo, aparente y común", que, según él, es "alguna medida de duración sensible

[1] No es preciso que nos ocupemos aquí de una dificultad especial de la teoría de Leibniz. A causa de su lógica de sujeto-predicado él no es capaz de mantener consistentemente que existen relaciones en absoluto. De este modo si a y b son mónadas, no tenemos $R(a, b)$, donde R es una relación, sino más bien $R_b(a)$ y $R_a(b)$, siendo R_a y R_b propiedades relacionales. En último lugar, como Bertrand Russell ha argumentado [1], la negativa de Leibniz a admitir relaciones le lleva a inconsistencia.

y externa (ya bien exacta o insuficiente) mediante el movimiento, que se usa comúnmente en lugar del tiempo verdadero". De modo semejante afirmaba que "el espacio absoluto, por su propia naturaleza, sin relación a ninguna cosa externa, se mantiene siempre semejante e inamovible", pero que, por otra parte, "el espacio relativo es alguna dimensión movible o medida de los espacios absolutos, que nuestros sentidos determinan por su posición respecto a los cuerpos, y que se toma, comúnmente, como espacio inamovible".

Todo esto es expuesto en un lenguaje bastante metafísico, y parte de lo que Newton mantuvo es, sin duda, realmente metafísico. Con todo, su problema era esencialmente matemático y puede ser establecido neutralmente como entre teorías absolutas y relativas del espacio y el tiempo. Por ejemplo, parte de lo que Newton quiso significar con el espacio absoluto fluyendo uniformemente, en contraste con el tiempo relativo, era simplemente lo siguiente: Si utilizamos como medida del tiempo la rotación aparente de las estrellas fijas alrededor de la Tierra, debido a que la Tierra está girando sobre su eje, nos encontramos con que las predicciones de la mecánica newtoniana de las posiciones y movimientos de los planetas no se ajustan con exactitud a las observaciones. Sin embargo, si se presupone que la rotación de la Tierra es ligeramente irregular, podemos cambiarnos a una nueva medida de tiempo, t', que está relacionada con la anterior t, mediante una función $t'=f(t)$. (De un modo sólo aproximado, $t'=t$.) Si la nueva escala de tiempo t' se aplica a las ecuaciones de la mecánica newtoniana, nos encontramos con que las predicciones de las posiciones y los movimientos de los planetas se vuelven exactas. En la terminología de Newton diríamos que t' es una medida de tiempo absoluto, mientras que t es una medida de tiempo relativo. Puede verse que no se trata de una cuestión metafísica, sino simplemente de la búsqueda de una función adecuada $f(t)$. No hace falta que consideremos a t' definida en términos de rotación en el espacio absoluto, ya que puede ser definida, y en la práctica tiene que ser definida, mediante la función $f(t)$, de la cual el argumento t se define mediante las relaciones con las estrellas fijas.

La misma lección, ilustrada aquí con relación al tiempo, puede ser aprendida, con relación al espacio, mediante el experimento de Newton del cubo en rotación, que se discute, también en la "Glosa a las Definiciones" de los *Principia*. Supóngase que se cuelga un cubo de una cuerda retorcida, de modo que conforme la cuerda se pone derecha se mueve en rotación rápidamente. También el agua del cubo, en virtud de la viscosidad, empezará a girar, y su superficie se hará cóncava. Finalmente, esta concavidad alcanzará un punto máximo cuando el agua del cubo esté girando a la misma velocidad angular que el propio cubo. Con todo, a partir de la concavidad del agua y de las leyes de movimiento y gravitación podemos deducir que el agua debe tener una considerable velocidad angular (igual a la del cubo). ¿En relación con qué es esta velocidad angular una velocidad angular? La respuesta de Newton es: "el espacio absoluto".

Otra respuesta es igualmente consistente, sin embargo, con los principios de la mecánica de Newton. ¿Por qué no podía haber afirmado Newton que la rotación era relativa a un conjunto de ejes determinado por la distribución general de la materia en el universo? Para Newton esta determinación de ejes habría tenido lugar mediante la referencia al sistema de las estrellas fijas; hoy en día podríamos hacerlo mediante referencia a las galaxias. Tal sistema sería adecuado para expresar la mecánica de Newton, sin necesidad de postular fuerzas adicionales; fuerzas centrífugas, por ejemplo. Tal conjunto de ejes se denomina sistema inerte.

Podría preguntarse, sin embargo, qué habría que decir si el experimento del cubo de Newton fuese realizado por físicos que viviesen en un planeta que, al igual que Venus, estuviese rodeado de nubes densas, más allá de las cuales no penetrasen jamás, de modo que no tuviesen conocimiento de cuerpos, tales como las estrellas fijas, externos a su planeta. En este caso podrían especificar, en principio, un sistema inerte mediante la relación a la superficie de su planeta. Esto sería, únicamente, un recurso transitorio, como resultaría manifiesto si mejorasen sus técnicas físicas de modo que en lugar del burdo experimento del cubo pudiesen llevar a cabo el experimento del péndulo de Foucault. En este experimento se permite que un péndulo se mueva en un plano vertical a partir de un fulcro que es capaz de girar sobre un eje vertical aproximándose todo lo posible sin llegar a la fricción. Cuando este experimento se realiza en la Tierra se observa que el plano de oscilación del péndulo gira una vez cada veinticuatro horas. Los físicos del planeta cubierto de nubes, de modo semejante, serían capaces de deducir del experimento del péndulo de Foucault que se encontraban en un planeta giratorio. Ahora bien, ¿cuáles dirían que son los ejes en relación a los cuales su planeta gira? No pueden referir la rotación a la superficie del propio planeta, como podrían haber hecho con relación al experimento del cubo. Comprenderían que cometerían un error si consideraban ejes que estuviesen fijos con respecto a la superficie del planeta como determinantes de un sistema inerte. (El experimento del cubo, a diferencia del experimento del péndulo de Foucault, habría sido, sencillamente, demasiado inexacto para mostrarles su equivocación.) Por tanto, quizá se inclinarían a decir que los ejes que habían descubierto como determinantes de un sistema inerte, y con relación a los cuales el plano de oscilación del péndulo de Foucault no giraría, son ejes fijos en el espacio absoluto.

Con todo, *no haría falta* que estos físicos dijesen tal cosa. ¿Por qué no podían decir, simplemente, que un sistema inerte estaba determinado por un conjunto de ejes que giraban con relación a su planeta? Podrían haber conservado incluso el conjunto de ejes que estaban fijos con relación al planeta y haber modificado adecuadamente las leyes de la mecánica. En el estado en que se encontraban, agobiados por las nubes, esto habría parecido probablemente una medida razonable. No resulta razonable, desde nuestro punto de vista actual, ya que determinaría, de modo arbitrario, un punto del espacio como importante cósmicamente (centro

del planeta), y, si todas las cosas del universo estuviesen girando alrededor de un punto, los objetos muy distantes tendrían que tener una velocidad mayor que la luz. (Volveremos a algunas de estas consideraciones en las págs. 180-2 más adelante.) Pero estas dificultades no aparecerían de modo manifiesto a nuestros astrónomos rodeados de nubes, ya que si su planeta era el único lugar del universo conocido por ellos, parecería razonable del todo que el centro del planeta fuese de importancia cósmica y que hubiese una ley de fuerzas centrípetas dependientes de él. (Compárese con la física aristotélica en la cual el centro de la Tierra forba parte de las leyes de la naturaleza.)

También sería posible que los físicos del planeta nublado tomasen otras medidas. Podrían postular que los sistemas inertes dependen de la distribución general de la materia en el universo, y a partir de aquí podrían deducir que de poder colocarse por encima de las nubes obtendrían testimonios visuales de tal sistema, con respecto al cual su planeta se vería que daba vueltas con una velocidad angular determinada.

Puede verse, por consiguiente, que habría varias alternativas posibles para los físicos del planeta nublado, con las que podrían evitar perfectamente una teoría absoluta del espacio. Al igual que el cubo de Newton, el péndulo de Foucault, proporciona testimonios que son del todo neutrales a la hora de elegir entre una teoría absoluta o relacional del espacio. Ernest Mach [3], que discutió el cubo de Newton de una forma muy parecida a aquélla en la que yo lo he hecho en las páginas precedentes, concluía que las afirmaciones acerca de los movimientos de un cuerpo "en el espacio" debían hacerse de modo que contuviesen una referencia no al espacio absoluto sino a todo el universo material. El principio de que la inercia de un cuerpo está determinada por la distribución de la materia en el universo ha sido, por consiguiente, denominado por Einstein "Principio de Mach" y es un punto de debate dentro de la moderna cosmología, ya que no es fácil, en modo alguno, modificar las ecuaciones de la teoría general de la relatividad, de forma que el principio de Mach sea obedecido. (Véanse págs. 180-2 más adelante.)

De hecho, el concepto de la distribución general de la materia en el universo no es lo bastante preciso para que podamos determinar un sistema inerte con suficiente exactitud dentro de la mecánica celeste (la teoría de los movimientos del Sol, los planetas, y otros miembros del sistema solar) que fue el objetivo primordial de Newton. Supóngase que poseemos un número de partículas P_1, P_2, ..., P_n. (El Sol, los planetas, la Luna, etc., no son partículas, sino que, puesto que son esféricos, es una consecuencia de la ley de la gravitación de Newton que se comporten gravitatoriamente como si toda su masa estuviese concentrada en su centro.) Entonces, nuestro problema consiste en asignar números m_1, m_2, ..., m_n a estas partículas (véase también la pág. 128, anteriormente), y un conjunto adecuado de ejes tal que si f_{kl} es el componente de la aceleración de P_k hacia P_l, tenemos en todos los casos $m_k f_{kl} = m_l f_{lk}$. (Los números m_1, m_2, ..., m_n, miden, por supuesto, las masas de las partículas.) Este pro-

blema no tiene solución analítica, pero puede solucionarse mediante métodos aproximativos. El conjunto de ejes así determinado es un sistema inerte. Ahora bien, si se le hubiese preguntado a Newton si su noción del espacio absoluto era realmente necesaria, o si un sistema inerte definido relacionalmente, tal como acabamos de hacer, era lo que él en realidad pedía, es muy posible que hubiera contestado que esto era esencialmente lo que él pensaba. En cualquier caso, probablemente habría admitido que, aparte de las consideraciones teológicas y metafísicas, todo lo que la *física* requiere es un sistema inerte que pueda ser definido relacionalmente.

Aun cuando la mecánica de Newton es del todo compatible con una teoría relacional del espacio y el tiempo, también es compatible con una teoría absoluta. Sin embargo, si queremos seguir las palabras del propio Newton literalmente y mantener que el espacio y el tiempo son absolutos, tenemos que aceptar una inconveniente consecuencia, a saber, la de que no podemos saber nunca *qué* sistema inerte corresponde a un conjunto de ejes en descanso en el espacio absoluto. Ello se debe a que existe una infinidad de sistemas inertes. Si S es un sistema de ejes inertes, entonces cualquier otro sistema S' que se mueva con velocidad uniforme con relación a S, también será un sistema inerte. Todo lo que pueden determinarnos los procedimientos del tipo descrito en el parágrafo anterior, sería las *aceleraciones* absolutas. Ninguna *velocidad* con respecto al espacio absoluto podría ser determinada mediante observaciones dentro de la mecánica. A partir de la teoría especial de la relatividad ya sabemos también que una velocidad, con respecto a un espacio absoluto, no puede ser determinada por medio alguno, como, por ejemplo, mediante experimentos ópticos. Un filósofo "verificacionalista" (verificationist)[2] u operacionalista, alegaría que carece de significado un discurso acerca del espacio absoluto. Con todo, si no somos "verificacionalistas" con relación al significado (y no debemos olvidar la incapacidad de los positivistas lógicos para proporcionar un criterio verificacionalista del significado adecuado), entonces podemos permitir que la afirmación de Newton de la existencia del espacio absoluto sea tanto significativa como consistente con sus presupuestos. Sin embargo, podemos seguir considerando inútil la afirmación de la existencia del espacio absoluto, como afirmación que no proporciona ninguna hipótesis científica interesante (debido a que no es comprobable).

Se podría interpretar la crítica de Leibniz a la teoría del espacio y el tiempo absolutos de Newton según el espíritu del parágrafo previo. Leibniz apelaba al "principio de la razón suficiente", que, libre de los matices teo-

[2] Por verificacionalista se entiende, *a grosso modo,* alguien que mantiene que el significado de una sentencia es una función del método de verificar la proposición que expresa. Si no existen medios de verificación, entonces la sentencia carece de significado. Aquí puede observarse una incoherencia, ya que si la sentencia carece de significado, y no expresa una proposición, ¿qué es lo que no se puede verificar? (Ryle ha criticado la teoría de la verificación del significado en este sentido [4].) La dificultad es pasada por alto por los defensores de la teoría de la verificación utilizando la palabra "proposición" ambiguamente significando a la vez "sentencia" y "lo que expresa una sentencia".

lógicos con que aparece en Leibniz, podríamos interpretar como "principio de la comprobabilidad". Leibniz señala que en la teoría absoluta del tiempo podríamos preguntar por qué Dios no creó todas las cosas un año antes de lo que lo hizo. Esto da lugar a una dificultad, ya que Dios no podía tener razón alguna para crear el universo en un instante del tiempo absoluto en lugar de otro, y es de suponer que todo lo que Dios hizo fue hecho por alguna razón. (Aun cuando supongamos que el mundo existe desde un pasado infinito, podríamos seguir preguntando por qué Dios creó el mundo tal cual es, en lugar de haber hecho las cosas un año antes o un año después.) No habría ninguna diferencia observable. Igualmente, respecto al espacio, podríamos preguntar por qué está todo donde está, en el espacio absoluto, en lugar de que todas las cosas estuviesen una yarda a la derecha, por ejemplo, de donde en realidad están. Leibniz piensa que ninguna cosa del mundo (o ninguna cosa que Dios haga) es arbitraria, o que no se deba a una razón. El principio de razón suficiente no parece ser analítico (verdadero por definición), pero, posiblemente, puede ser interpretado como una anticipación de la objeción verificacionalista a la falta de significado de las nociones de espacio y tiempo absolutos. He preferido reinterpretar la forma teológica del principio de Leibniz de modo diferente, como un principio de comprobabilidad, no de significado.

Podemos concluir, cuando menos, que en el contexto de la mecánica newtoniana la teoría del espacio y el tiempo absolutos es científicamente ociosa. Newton puede haber dicho todo lo que tenía que decir en realidad, desde un punto de vista científico, hablando de la necesidad de determinar sistemas inertes, y los sistemas inertes pueden discutirse consistentemente dentro de una teoría relacional del espacio y el tiempo. Esto no significa, sin embargo, que la teoría del espacio-tiempo absoluto (no podemos seguir ya hablando del espacio y tiempo separadamente) sea necesariamente ociosa en la física moderna. Veremos que no le es fácil a la teoría *general* de la relatividad prescindir del espacio-tiempo absoluto, y el cosmólogo J. A. Wheeler ha aspirado, incluso, a hacer que el espacio-tiempo fuese toda la materia del mundo, siguiendo la tradición de Descartes. Las ideas de la matemática moderna permiten que una posición esencialmente cartesiana sea desarrollada de tal modo que al menos no está expuesta a las simples objeciones de la posición del propio Descartes. Es cierto que dentro de la teoría *especial* de la relatividad una teoría relacional del espacio-tiempo parece del todo plausible, aunque aquí incluso no es necesaria. Sí tenemos, desde luego, que abandonar una teoría absoluta del espacio y el tiempo considerados separadamente, pero esto no significa que nos veamos abocados a una teoría relacional del espacio-tiempo, como un todo. Desde luego que, si no vamos más allá de la teoría *especial,* tal postulación del espacio-tiempo absoluto no tiene sentido, al igual que carece de sentido la teoría newtoniana del espacio y tiempo absolutos, pero no es inconsistente con la teoría. Una breve discusión del espacio-tiempo tal como aparece dentro de la teoría especial de la relati-

vidad no sólo será útil en sí misma si no que servirá de preparación para discutir la pugna entre la teoría absoluta y la teoría relacional del espacio-tiempo dentro de la cosmología moderna. Vale la pena añadir, además, que en el estado actual de conocimientos la relatividad especial es filosóficamente importante por propio derecho: no debe ser considerada como una teoría incorrecta que proporciona una aproximación de las predicciones de la teoría verdadera, que fue la línea que seguimos (en el capítulo 3) con la termodinámica fenomenológica en su relación con la termodinámica estadística. De hecho la relatividad *general* no es, en modo alguno, una teoría bien establecida, y sólo sirve para discutir problemas gravitacionales. Es la teoría especial la que se utiliza ampliamente, como, por ejemplo, en la microfísica.

El ejemplo kantiano de la mano izquierda y la mano derecha

En sus *Prolegomena to Any Future Metaphysics* ("Prolegómenos a cualquier futura Metafísica", secc. 13) [5] Kant alude al caso de cosas tales como la mano izquierda y la mano derecha. Supóngase que el universo contuviese un único objeto, una mano izquierda. Todas las relaciones entre sus partes serían las mismas que lo habrían sido si hubiera habido una mano derecha. Por tanto, la teoría relacional del espacio no puede hacer justicia a las diferencias entre una mano izquierda y una mano derecha. El relacionista, sin embargo, puede dar, fácilmente, una respuesta. Puede replicar que de haber un único objeto, una mano, en el universo, carecería de significado denominarle ya bien mano izquierda o ya bien mano derecha. O, por lo menos, el relacionista diría que esto es lo que se afirmaría de acuerdo con el presupuesto de que las leyes de la naturaleza son simétricas con relación a la imagen de un espejo. El reciente descubrimiento en física de la no conservación de la paridad sugiere que, siempre que sepamos si la mano consiste en materia o anti-materia, habrá significado en afirmar que era una mano izquierda o una mano derecha. Pero esto, una vez más, consistiría en una *relación* entre "la facultad de ser mano" (handedness) de las partículas elementales de las cuales estaba compuesta la mano y la "facultad de ser mano" de la mano como un todo. Una vez más, no habría necesidad de espacio absoluto. Por lo demás, habría dificultades en la significatividad de afirmar si la mano consistía en materia o antimateria, de ser el único objeto del universo. Denominemos a la materia "la clase más común de partículas del universo" (o de nuestra parte del universo). De este modo los electrones son más comunes que los positrones. Entonces, por definición, nuestro universo de una sola mano consistiría en materia, no en antimateria.

El espacio-tiempo de la relatividad especial

Considérese el conjunto de ejes $O'x'$, $O'y'$, $O'z'$ de la figura 4 que se mueven con velocidad constante v en la dirección Ox de un conjunto paralelo de ejes Ox, Oy, Oz. De acuerdo con la transformación prerrelativista o "galileana" las nuevas coordenadas están relacionadas con las antiguas mediante las ecuaciones

$$x' = x - vt$$
$$y' = y$$
$$z' = z$$

a lo cual podemos añadir, para lograr la completitud (completeness)

$$t' = t$$

Si aplicamos las transformación anterior $(x, y, z, t) \rightarrow (x', y', z', t')$ a las ecuaciones electromagnéticas de Maxwell, nos encontramos con que estas ecuaciones cambian de forma matemática y, además, que implican (como el sentido común nos haría esperar) que la velocidad de la luz en

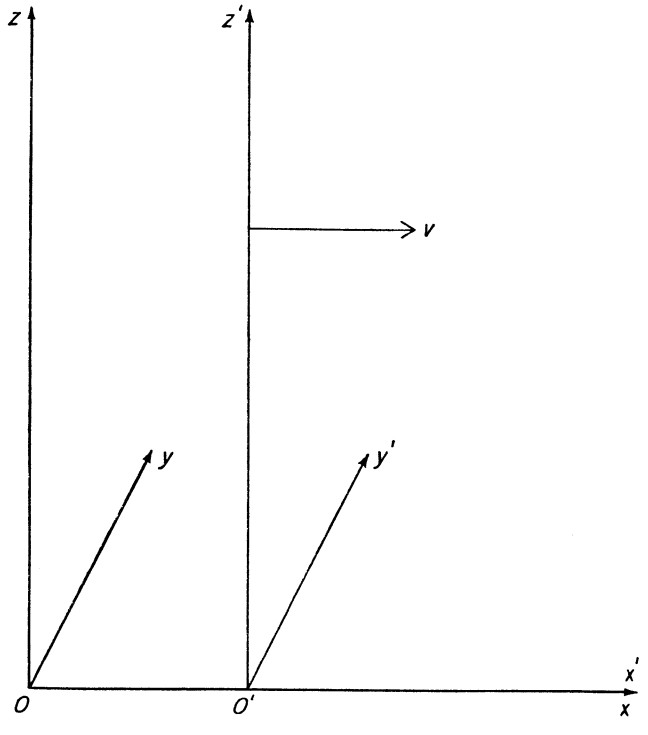

Figura 4.

relación con un sistema de ejes, se diferenciará de la que haga relación al otro sistema mediante v. El resultado de los experimentos de Michelson-Morley, y otros experimentos ópticos, sugerían, por el contrario, que cualquiera que fuese el sistema de ejes sería descubierto que la velocidad de la luz era exactamente la misma por los experimentadores que estén en descanso con relación a estos ejes. Este hecho, sorprendente de acuerdo con nuestras ideas vulgares, es explicado por la teoría especial de la relatividad de Einstein.

De acuerdo con esta teoría la transformación galileana es sustituida por lo siguiente (suponiendo, una vez más, que la velocidad v de uno de los sistemas con respecto al otro es paralela a Ox y siendo c la velocidad de la luz):

$$x' = \frac{x - vt}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$y' = y$$

$$z' = z$$

$$t' = \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Obsérvese cuanto más simétricas parecen estas ecuaciones si elegimos unidades de espacio y tiempo de modo que $c=1$. Obsérvese también que estas ecuaciones se aproximan a las galileanas cuando v es pequeña.

Cuando se utilizan las transformaciones de la relatividad especial en lugar de las galileanas, las ecuaciones del electromagnetismo de Maxwell conservan su forma, e implican el mismo valor para la velocidad de las ondas electromagnéticas (luz) cualquiera que sea el sistema de ejes. El resultado negativo del experimento de Michelson-Morley se sigue como consecuencia, como también los resultados de otros experimentos ópticos (Ives-Stilwell y Kennedy-Thorndike) que son incompatibles con las ideas pre-relativistas. De modo semejante otros resultados experimentales de la óptica, tales como el desvío y el efecto de Doppler se pueden deducir de la teoría de Maxwell junto con las nuevas transformaciones.

H. A. Lorentz se había propuesto explicar el resultado negativo del experimento de Michelson-Morley mediante una teoría que implicaba que los cuerpos en movimento se contraen en la dirección del movimiento. Es decir, la velocidad de la luz varía, en realidad, de un sistema a otro, pero este cambio es imperceptible debido a una contracción exactamente compensadora de nuestros instrumentos de laboratorio en su dirección de movimiento. Aunque esta teoría pueda parecer un tanto *ad hoc,* no es, en realidad, tan poco plausible, considerando que la materia consiste en gran medida en partículas cargadas de electricidad (y en aquellos mo-

mentos se creía que consistía enteramente en partículas cargadas de electricidad), y así Lorentz pudo argumentar con cierto éxito, que las mismas leyes que llevaban a un cambio en la velocidad de la luz, tal como resultaba medida en un sistema u otro, llevarían también a una contracción de cuerpos materiales que harían que el cambio de la velocidad fuese imperceptible. Así pues, aunque la hipótesis de Lorentz fuese inverificable, recibiría apoyo indirecto de las consideraciones teóricas mediante las cuales él se esforzó en justificarla. Resulta, sin embargo, interesante lo señalado por Adolf Grünbaum [6] con respecto a que la hipótesis de Lorentz no es inverificable y, por tanto, no es *lógicamente ad hoc,* aunque, sin embargo, debido a que Lorentz no conocía estas pruebas de comprobación puede decirse que es *psicológicamente ad hoc.* De hecho la hipótesis de la contracción (éter-teórica) de Lorentz puede ser refutada por el experimento de Kennedy-Thorndike. Este es semejante al experimento más famoso de Michelson-Morley, excepto en que, en lugar de tener los dos brazos interferómetros (que están colocados formando un ángulo recto) de un largo igual, son muy distintos entre sí.

Volvamos a las denominadas transformaciones de Lorentz, tal como aparecen en la teoría especial de la relatividad (las transformaciones de la pág. 167 anterior). Aquí son simplemente una transformación matemática de un sistema de coordinadas a otro y no representan verdaderas contracciones (causadas por el movimiento) como en la teoría de Lorentz. Como fue descubierto por Minkowski, las transformaciones de Einstein representan simplemente una rotación de ejes en el espacio-tiempo. Si troceamos un pastel cilíndrico en distintos ángulos obtenemos distintas secciones transversales del pastel, algunas más elípticas o alargadas que otras, pero, desde luego, no hay ninguna expansión física ni contracción del pastel. Algo bastante parecido a esto ocurre con la relatividad especial: un estado instantáneo de una regla en reposo en relación a nosotros corresponde a un corte recto del objeto cuatridimensional que es la regla a través de toda su historia, mientras que si nos estamos moviendo en relación a la regla, lo que vemos en un instante corresponde a un corte oblicuo. Hay una diferencia con relación al caso del pastel cilíndrico, que consiste en que los cortes oblicuos son más alargados que los rectos, ya que debido a la geometría especial del espacio-tiempo los cortes oblicuos de la regla son *más cortos* que los cortes rectos. De aquí la denominada "contracción".

Las ecuaciones del movimiento de Newton son constantes con relación a la transformación galilea, pero las ecuaciones del electromagnetismo no son constantes con respecto a esta transformación. Por otra parte, mientras que las transformaciones de Lorentz de la relatividad especial dejan a las ecuaciones del electromagnetismo constante, no dejan a las leyes de la mecánica newtoniana constantes. Para alcanzar esta constancia Einstein tuvo que modificar las leyes de la mecánica. En el caso de los cuerpos que se mueven a velocidades reducidas con relación a nuestro sistema de ejes, las nuevas ecuaciones predicen los mismos resultados,

dentro de los límites del error observacional, como ocurre con las newtonianas, pero llevan a resultados comprobablemente distintos para velocidades elevadas, y las ecuaciones modificadas han sido confirmadas mediante observaciones de partículas aceleradas. Es importante resaltar que la comprobabilidad de la relatividad especial depende en gran medida de estas predicciones mecánicas, mientras que los críticos de la relatividad se contentan a veces con intentar explicar los experimentos ópticos de un modo no relativista. La teoría especial de la relatividad se comprueba de muchos modos, y sus predicciones tienen que ser tomadas en cuenta incluso en la ingeniería (en el diseño de ciclotones, por ejemplo) y no se basa simplemente en unos cuantos experimentos ópticos.

Debido a que la ecuación $t'=t$ de la transformación galileana es sustituida en la relatividad especial por la ecuación

$$t'=\frac{t-\frac{vx}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}$$

puede verse que el que los acontecimientos distantes sean o no simultáneos depende en parte de v. Es decir, si los acontecimientos distantes son simultáneos en un sistema para el cual $v=0$, entonces, en un sistema que se mueva con velocidad v en relación con este sistema, uno de los acontecimientos será anterior o posterior que el otro. Cuando $x=0$, por supuesto, los acontecimientos que sean simultáneos de acuerdo con la transformación galileana serán simultáneos de acuerdo con la de Lorentz.

En el laboratorio operamos con cosas que están aproximadamente en el mismo lugar, y por tanto las medidas directas de laboratorio de la simultaneidad no se ven afectadas por el cambio de las ideas newtonianas a las relativistas.

Consideremos ahora el hecho, demostrado por Minkowski, de que las transformaciones de Lorentz pueden ser consideradas simplemente como una rotación de ejes en el espacio-tiempo. Tenemos que presuponer que la métrica del espacio-tiempo es dada no por

$$ds^2=dx^2+dy^2+dz^2+dt^2$$

como ocurriría si fuese euclideana, si no por la métrica semi-euclidiana [3]

$$ds^2=dt^2-dx^2-dy^2-dz^2$$

(Por razones de conveniencia he supuesto que hemos elegido unidades

[3] Es semi-*euclidiana* porque una sección tri-dimensional en ángulo recto con el eje del tiempo es euclidiana. Es *semi*-euclidiana a causa del signo menos: las secciones del otro lado (a lo largo del eje temporal) son hiperbólicas.

de tiempo y espacio de tal modo que la velocidad de la luz es igual a uno. De lo contrario tendríamos en las dos ecuaciones anteriores c^2t^2 en lugar de t^2.)

Considérese ahora la figura 5, en la cual se muestra la dirección del tiempo *Ot* y la dirección del espacio *Ox*. (Por supuesto sólo puedo dibujar dos dimensiones en un trozo de papel.)

Hagamos que una partícula se mueva con velocidad constante *v* relativa a una partícula en reposo en *O*. Hagamos que todas las posiciones sucesivas de espacio-tiempo estén en la línea *Ot'*. Esta línea se denomina la línea del mundo de la partícula y, evidentemente, la partícula, consi-

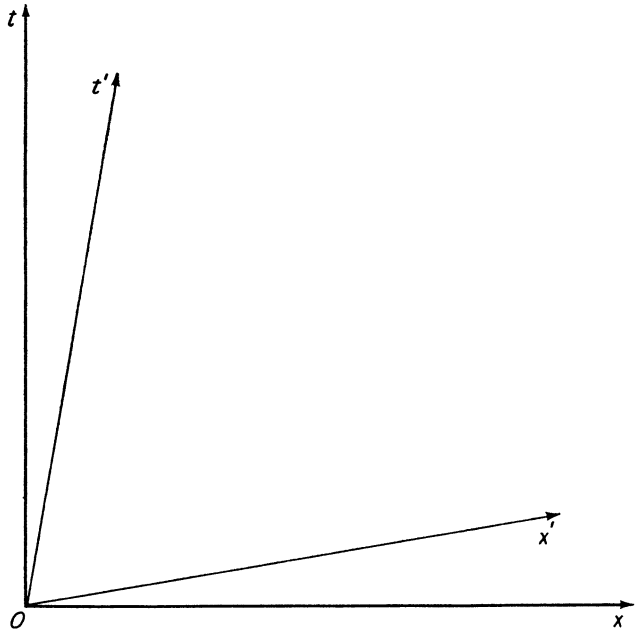

FIGURA 5.

derada como un cuerpo espacio-temporal de cuatro dimensiones, está a lo largo de *Ot'*. Evidentemente también *Ot* es la línea del mundo de una partícula en reposo en el sistema de ejes original. Es natural, por consiguiente, que consideremos a *Ot'* como el eje temporal de un sistema de ejes con relación al cual el cuerpo en movimiento se diría que (considerándolo como una entidad tridimensional permanente) "está en reposo". Ahora si hacemos girar a *Ox* cuando *Ot* gira a *Ot'* viene a ponerse a lo largo de *Ox'*. Esto parece un poco contra-intuitivo en el diagrama, ya que *Ox'* va *por encima* de *Ox*, no por debajo como cabría esperar. Esto se debe a la geometría semi-euclidiana del espacio-tiempo, que hace que los diagramas dibujados en trozos de papel euclidianos parezcan bastante raros. Obsérvese que la velocidad *v* es simplemente la tangente del ángulo *tOt'*. Es un simple ejercicio de la geometría de coordenadas el descubrir las

ecuaciones para transformar el sistema de coordenadas tOx en el $t'Ox'$. Resultan ser las mismas que las de la transformación de Lorentz.

La teoría especial de la relatividad resulta mucho más comprensible que antes cuando la contemplamos a la luz de la geometría del espacio-tiempo. Desde luego resulta mucho más atractiva que la antigua teoría newtoniana. Ya que, ¿qué ocurre en la teoría newtoniana si, como parece plausible, continuamos considerando nuestro eje temporal como una línea a lo largo de la cual coincide la línea del mundo de una partícula en reposo de nuestro sistema? En la teoría newtoniana entonces hacemos girar a Ot pero no hacemos girar a Ox. Hay, por consiguiente, como Minkowski ha señalado [7], una cierta elegancia matemática en las transformaciones de Lorentz, de la que carecen las newtonianas.

Quizá sea preciso señalar que la asimetría en la apariencia del diagrama es desorientadora. Se debe, una vez más, a tratar de representar un espacio no euclidiano en un trozo de papel euclidiano. En nuestro papel el sistema de coordenadas rectangular de $t'Ox'$ parecerá romboide. Pero igualmente podríamos haber dibujado $t'Ox'$ formando un ángulo recto y a tOx como oblicuo, como en la figura 6.

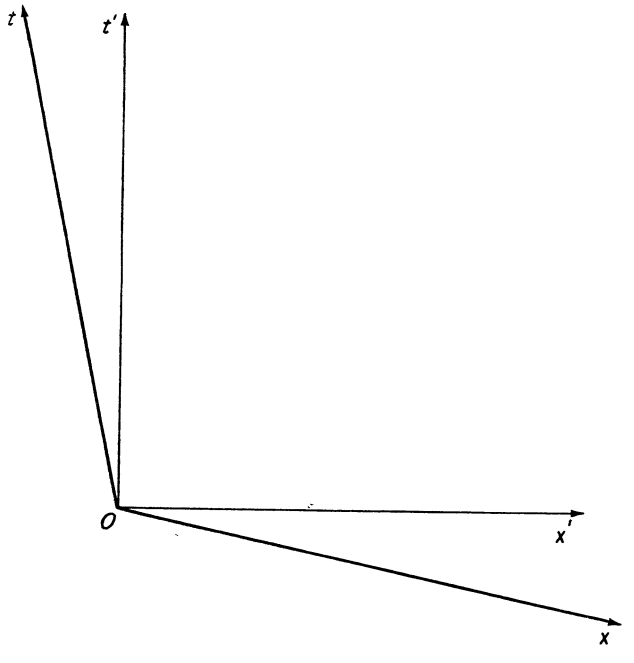

FIGURA 6.

De acuerdo con las elocuentes palabras de Minkowski: "Por consiguiente, el espacio por sí mismo, y el tiempo por sí mismo, se considera que se desvanecen en meras sombras, y sólo un tipo de unión de ambos conservará una realidad independiente" [8].

Es importante aprender a pensar en términos de una imagen espacio-temporal cuatridimensional del mundo, puesto que Minkowski ha demostrado la indisoluble unión entre el espacio y el tiempo. Hay que evitar un traspiés. Es el que dan, con frecuencia, los físicos, especialmente en obras escritas para el público en general. Puesto que el tiempo ya está implicado en el espacio-tiempo, no podemos hablar de cambio en el espacio-tiempo en ningún otro sentido si no en el de que algunas partes del espacio-tiempo sean distintas de las otras. En esta última sentencia "sean" no es, desde luego, un auténtico tiempo verbal presente, si no que se trata del "presente" sin tiempo verbal que tenemos en matemáticas, como cuando decimos que π es un número irracional. Cuando decimos que π es un número irracional no queremos decir que sea ahora un número irracional. En lugar de decir que un cuerpo pasa de ser rojo a ser verde deberíamos decir que una sección transversal espacial anterior del cuerpo espacio-tiempo (de cuatro dimensiones) es (sin tiempo verbal) roja y que una posterior es (sin tiempo verbal) verde. También es evidentemente ilegítimo hablar de un cuerpo o de una señal moviéndose a través del espacio-tiempo. El concepto de movimiento se sustituye ahora por la noción de las inclinaciones relativas de las líneas del mundo. Así en la figura 7 las partículas *AB* y *CD* están más lejanas entre sí en t_2 que en t_1. De modo semejante, en vez de decir que el movimiento relativo de la segunda con respecto a la primera es mayor en t_1 que en t_2, decimos que

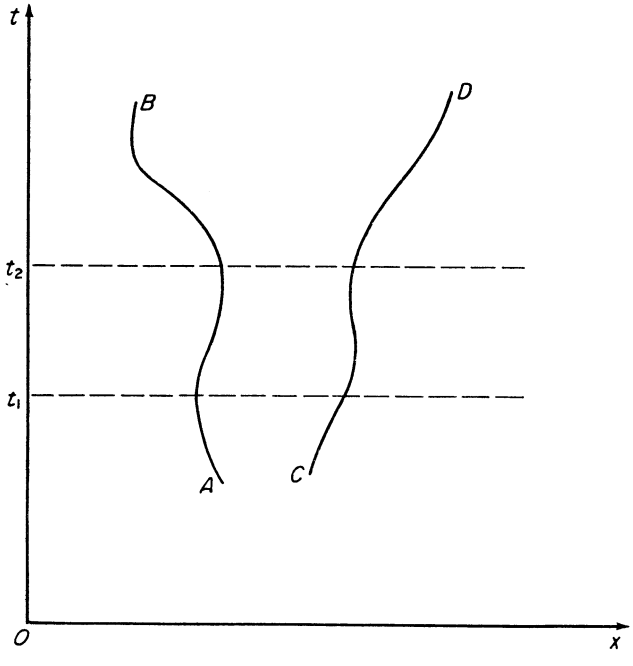

FIGURA 7.

la inclinación de una línea del mundo con respecto a la otra es mayor en t_1 que en t_2. No debemos hablar de cambios o movimientos en el sentido ordinario, ya que el tiempo está ya en nuestra representación [9].

Desde luego, sería desorientador afirmar que de acuerdo con la teoría de la relatividad el futuro está "ya en existencia". El futuro para mí, en el momento de escribir este libro, contiene fechas tales como el año 1970. Decir que los acontecimientos del año 1970 están ya en existencia sugiere que el año 1970 existe ahora, es decir, en 1967, lo cual es absurdo. Pensar de este modo es confundir el presente sin tiempo verbal de "está en el año 1970" con el tiempo presente "está ahora en 1970". El mundo de cuatro dimensiones contiene un lugar para todas las fechas posibles a lo largo del eje temporal, y, evidentemente, todas las cosas no son simultáneas. Ni tampoco hace este eje que ningún acontecimiento sea eterno: pensar tal cosa sería cometer el absurdo de tratar de hacer que un punto del mundo espacio-temporal cubriese una línea infinita que se extendiese del pasado al futuro. Por lo demás, el mundo de cuatro dimensiones no implica un determinismo. El determinismo sugiere que mediante un conocimiento completo de todas las cosas en t_1 (por ejemplo en la línea discontinua más baja de la figura 7, si podemos olvidar todas las dimensiones espaciales menos una), nosotros podemos, dado un conocimiento de todas las leyes de la naturaleza, deducir el estado completo del mundo en t_2 (por ejemplo, todos los acontecimientos a lo largo de la raya discontinua superior). No hay nada en el dibujo del espacio tiempo que implique esto o sea consistente con esto. La posición relativa a que el mundo es un mundo espacio-temporal es del todo neutral con respecto a la elección entre determinismo o indeterminismo.

La paradoja del reloj

Una discusión de la denominada paradoja del reloj de la relatividad especial ayudará a destacar el poder y atractivo intuitivo de la imagen del mundo de cuatro dimensiones. Desde luego, una vez que uno comienza a pensar cuatridimensionalmente, resulta difícil imaginar cómo nadie puede haber pensado jamás que hubiera nada realmente paradójico en la paradoja del reloj.

Una consecuencia de la ecuación de Lorentz

$$t' = \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

es la siguiente: Considérense dos relojes que están juntos en el punto O del espacio-tiempo. Uno está en reposo en el sistema (Ox, Ot): en otras

palabras, considerado como objeto cuatridimensional se extiende a lo largo de *Ot*. El otro se está moviendo a una velocidad *v*, y considerado como objeto cuatridimensional se extiende a lo largo de *Ot'*. Véase la figura 8. Considérese un acontecimiento en *T* de la historia del primer reloj. Evidentemente *T* está en *Ot*. Con referencia al sistema (*Ox, Ot*) el acontecimiento de la historia del segundo reloj que es simultáneo a *T* está en el punto *T'* de *Ot'*, tal que *TT'* es paralelo a *Ox*. Ahora bien, *OT'* es más corto que *OT*. (Parece más largo en el diagrama, que está dibujado en un trozo de papel con la métrica euclidiana dada para $ds^2=dt^2+dx^2$, mientras que en el espacio-tiempo tenemos $ds^2=dt^2-dx^2$.) Esto se expresa de un modo un tanto desorientador diciendo que los relojes en el sistema en moviminto corren despacio con relación al sistema en reposo. De modo semejante, con referencia al sistema (*Ox', Ot'*) el acontecimiento de *Ot* que es simultáneo con *T'* es *T''*, siendo *T''T'* paralelo a *Ox'*. *OT''* resulta ser más corto que *OT'*, exactamente igual que *OT'* era más corto que *OT*. Esto se expresa, nuevamente, de un modo desorientador diciendo que los relojes del sistema (*Ox, Ot*) corren despacio comparados con los relojes en el sistema (*Ox', Ot'*). Es fácil comprobar que no hay ninguna contradicción cuando se dice que los relojes del segundo sistema corren despacio comparados con los relojes del primer sistema, y que los relojes del primer sistema corren despacio comparados con los relojes del segundo sistema. No hay contradicción alguna en afirmar que $OT' < OT$ y que $OT'' < OT'$.

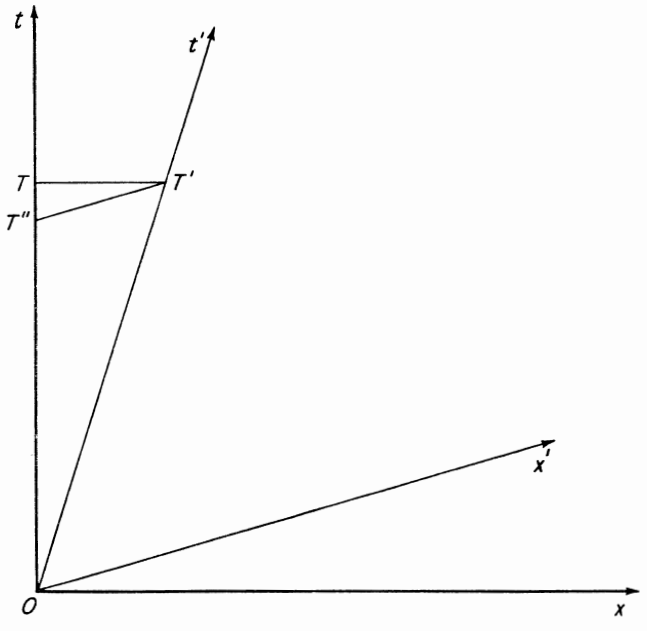

FIGURA 8.

De hecho, hagamos que T tenga las coordenadas (Ox, Ot) $(x_1, t_1 \cdot OT = t_1$ y $OT' = t'_1$. Entonces, de acuerdo con la cuarta ecuación de la transformación de Lorentz

$$t'_1 = \frac{t_1 - \frac{vx_1}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

siendo v la velocidad del segundo reloj con respecto al primero. Evidentemente $v = \frac{x_1}{t_1}$. Y, por tanto, tenemos

$$t' = \frac{t_1 - \frac{vx_1}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} = t_1 \sqrt{1 - \frac{v^2}{c^2}}$$

De modo semejante hagamos que T'' tenga la coordenada Ox', x'_2 (que será un número negativo). La coordenada Ot' de T'' es, evidentemente, t'_1. La velocidad del primer sistema con relación al segundo es $-v$ y, por tanto, tenemos que la coordenada Ot de T'' es t_2, siendo

$$t_2 = \frac{t'_1 - \frac{(-v)\, x_2}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{t'_1 + \frac{vx_2}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Pero $-v = \frac{x_2}{t'_1}$ y, por tanto,

$$t_2 = t'_1 \sqrt{1 - \frac{v^2}{c^2}}$$

lo cual confirma que tanto $OT' < OT$ como $OT'' < OT'$.

Ahora, con relación a la propia paradoja del reloj: supóngase que existen dos gemelos, Jack y Jim. Mientras que Jack permanece en la Tierra Jim se lanza en un cohete espacial a Alpha Centauri, que está a cuatro años luz de distancia, y regresa de nuevo. El cohete de Jim viaja a la mitad de la velocidad de la luz y (si podemos olvidarnos de las cuestiones de aceleración y contra-aceleración hacia y desde la velocidad del viaje, que es una cuestión que discutiré en seguida), el tiempo en el sistema de coordenadas (de la Tierra) de Jack, entre la partida y el regreso de Jim, es de dieciséis años. Sin embargo, el sistema de coordenadas de Jim

se está moviendo con una velocidad $\frac{c}{2}$ con relación al de Jack, y, por tanto, en el sistema de Jim el viaje al exterior lleva $8\sqrt{1-\frac{\frac{1}{4}c^2}{c^2}}=4\sqrt{3}$ años. En su viaje de regreso se mueve con una velocidad $-\frac{c}{2}$ con relación a Jack, y, por tanto, en su sistema de coordenadas el viaje de regreso también lleva $4\sqrt{3}$ años. Por tanto, su viaje de ida y vuelta lleva $8\sqrt{3}$ años. Jim es, por consiguiente, algo más de dos años más joven que Jack cuando regresa.

La paradoja del reloj procede del siguiente razonamiento falaz. En nuestros cálculos hemos supuesto que Jack estaba en reposo y Jim en movimiento con una velocidad de $+v$ o $-v$ con relación a él. Igualmente, se dice, podríamos considerar a Jim en reposo y que la velocidad de Jack con relación a Jim fuera también $\pm v$. Así pues, después del viaje Jack sería poco más de dos años más joven que Jim. Considerando las dos partes de la argumentación a la vez, tenemos que Jack es más joven que Jim y que Jim es más joven que Jack, lo cual es imposible.

La falacia del razonamiento consiste en que el primer cálculo (que mostraba a Jim más joven que a Jack) era correcto, debido a que Jack ha estado siempre en el mismo sistema inerte. Sin embargo Jim tuvo que ser acelerado y contra-acelerado a Alpha Centauri y, por tanto, habría tenido algunas experiencias fisiológicas muy dolorosas que Jack no ha tenido) y, por ello, la segunda parte de la argumentación no es válida. En la relatividad especial es permisible considerar a un conjunto de ejes en reposo sólo si son por completo un conjunto inerte de ejes.

Ahora podemos demostrar la validez de la imagen espacio-temporal del mundo de Minkowski. Si dibujamos las líneas del mundo de Jack y Jim el dibujo mostrará claramente la asimetría que indicábamos en el parágrafo previo, y la conservación de la juventud relativa por parte de Jim se considerará como una simple consecuencia de la geometría del espacio-tiempo.

En la figura 9, *Ot, Ox* son los ejes de Jack. Jack es un "gusano" cuatridimensional que se extiende a lo largo de *Ot*. *A* es la posición espacio-temporal de la vuelta de Jim alrededor de Alpha Centauri. Este abandona a Jack en *O* y vuelve a Jack en *B,* y, por ello, es un "gusano" espacio-temporal inclinado, extendido a lo largo de *OAB*. El resto del diagrama carece de importancia, pero he dibujado en él *Ot', O',* que son los ejes de Jim en el viaje al exterior y *Ot", Ox",* que son sus ejes en el viaje de regreso, a fin de mostrar que, como se observó previamente, no está en la misma situación inerte todo el tiempo. La asimetría, que la argumentación de la paradoja del reloj trata de negar, se ve muy claramente en la línea recta de Jack (en cuanto objeto cuatridimensional) y la inclinada de Jim. El tiempo de envejecimiento de Jack entre los acontecimientos *O* (marcha de Jim) y *B* (vuelta de Jim) es *OB*. El tiempo de envejecimiento

de Jim es la longitud de su línea de mundo, que es $OA+AB$. Intuitivamente $OA+AB$ parece mayor que OB, pero esto se debe al papel euclidiano en que está impreso este libro. En la geometría del espacio-tiempo $OA+AB$ es menor que OB. Pero admitido el carácter euclidiano de nuestro papel, no hay nada sorprendente en la paradoja del reloj: la relativa juventud de Jim es simplemente debida a que dos lados de un triángulo suman una longitud distinta de la del tercer lado.

En la discusión anterior he omitido los períodos de aceleración y contra-aceleración. Sin embargo, esto no importa demasiado, ya que al elegir una estrella suficientemente lejana (quizá bastante más distante que Alpha Centauri, que era la de nuestro ejemplo) podemos asegurarnos de que los períodos de aceleración y contra-aceleración son todo lo pequeños

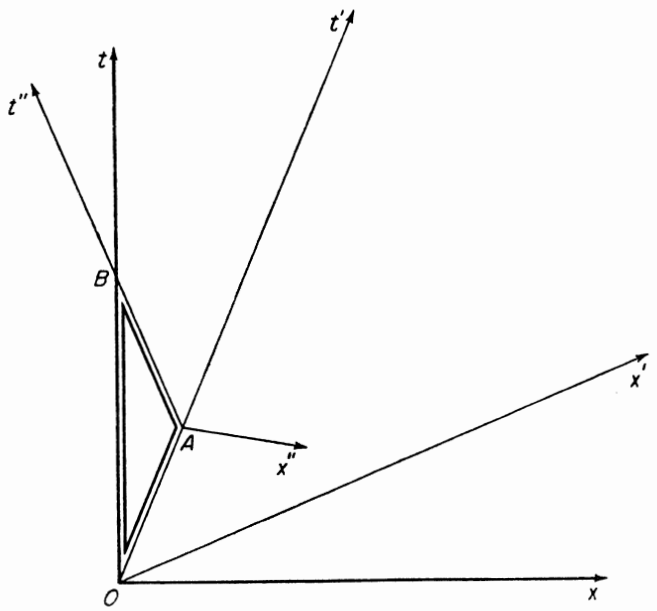

FIGURA 9.

que deseemos en comparación con el tiempo total. De este modo, si volvemos a dibujar el triángulo de nuestra figura de modo que muestre los períodos de aceleración y contra-aceleración, contendrá partes curvas como en la figura 10, pero éstas pueden ser muy pequeñas comparadas con los lados de un triángulo, y nuestro cálculo anterior puede ser aproximadamente correcto. En cualquier caso, un cálculo más complicado incluyendo un cómputo de la longitud de las partes curvas mostrará que el viaje de ida y vuelta es una línea del mundo más corta que la línea del mundo del gemelo que se quedó en casa.

En realidad podemos narrar una historia semejante con relación a Jack y a Jim, incluso si no presuponemos ninguna aceleración ni contra-

aceleración en absoluto. Supóngase que el universo fuese un universo finito (y sin expansión) de curvatura positiva constante. Supóngase que Jack se quedase en casa y que Jim estuviese en un cohete que pasase a Jack con una velocidad v en línea recta. Entonces, al igual que un marino navegando alrededor de la Tierra, Jim retornaría junto a Jack nuevamente. (¡Claro que en este caso tendríamos que suponer que Jack y Jim eran prácticamente inmortales!) Jim será más joven que Jack a su retorno, ya que si Jack ha envejecido un tiempo t entonces Jim habrá envejecido $t\sqrt{1-\frac{v^2}{c^2}}$. (Este caso fue mostrado y discutido por E. A. Mil-

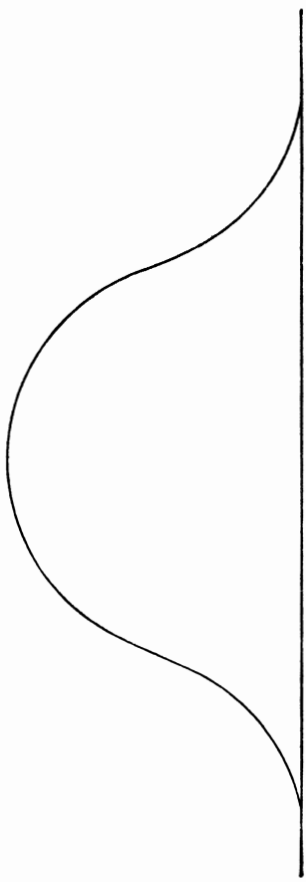

FIGURA 10.

ne y G. J. Whitrow [10].) Podemos demostrar esto en un diagrama del tipo de Minlowski si cambiamos el dibujo ligeramente. Dibújese el diagrama no en una hoja de papel plana, si no en un cilindro, como en la figura 11. Entonces la línea que va alrededor del cilindro es más corta

(aunque a causa de nuestras intuiciones euclidianas parezca más larga) que la línea recta *Ot* que es la línea del mundo de Jack. En el diagrama puede verse en qué sentido Jack está en reposo con relación al universo como un todo, mientras que Jim no lo está. La diferencia entre las líneas del mundo de Jack y de Jim es absoluta (topológicamente), Jim da la vuelta al cilindro mientras que Jack no lo hace. Aunque los cálculos sobre estas líneas utilizarían únicamente las transformaciones de Lorentz de la relatividad especial, nos estamos acercando más a las ideas cosmológicas asociadas con la teoría general de la relatividad, ya que es preciso que consideremos la estructura del espacio-tiempo desde un punto de vista cosmológico.

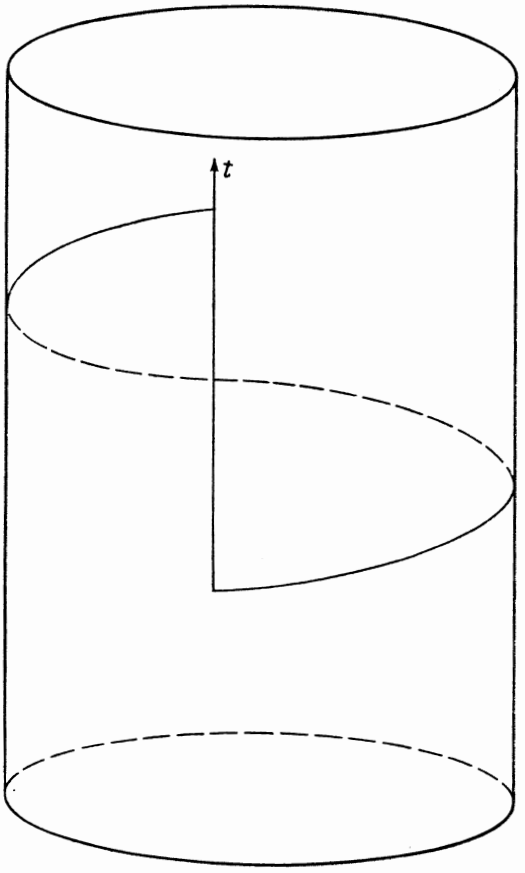

FIGURA 11.

Puede observarse que la teoría especial de la relatividad es inconsistente con la teoría de Newton de un tiempo y un espacio absolutos considerados separadamente. Por otra parte, parece ser perfectamente consistente con una teoría absoluta del *espacio-tiempo*. ¿Es el espacio-tiempo una plenitud absoluta o se trata simplemente de un sistema de relaciones

entre acontecimientos, ordenados cuatridimensionalmente? La teoría especial de la relatividad es del todo neutral en relación a la elección entre estas interpretaciones, exactamente igual que la mecánica de Newton, en oposición a su metafísica, es neutral respecto a la elección entre teorías absolutas y relativas del espacio. También se da, a su vez, la misma arbitrariedad. Al igual que no contamos con ninguna razón especial para suponer que ningún sistema inerte determinado esté en reposo en el espacio absoluto, y exactamente al igual que no tiene importancia ninguna para la física que supongamos que todas las cosas del universo estén una milla a la derecha o a la izquierda, delante o atrás, arriba o abajo, del mismo modo surge una situación idéntica en el mundo de Minkowski con relación al espacio-tiempo. No importaría en absoluto que todo lo que aparece en nuestro diagrama de Minkowski se empujase hacia arriba o hacia abajo, hacia la derecha o hacia la izquierda, o que nuestro diagrama girase a través de cierto ángulo. Todas las relaciones físicamente importantes permanecerían iguales. Así que, aunque el espacio-tiempo absoluto sea consistente con la relatividad especial, no desempeña ningún papel físicamente importante dentro de la teoría. En la teoría general de la relatividad, sin embargo, hay buenas razones para suponer que el espacio-tiempo absoluto puede desempeñar un papel esencial dentro de la teoría. En cualquier caso las hipótesis cosmológicas que incluyan la teoría general de la relatividad tendrán que hacer uso, probablemente, del espacio-tiempo absoluto, de un modo esencial.

El espacio-tiempo en la relatividad general

En la actualidad la cuestión de si la relatividad general puede ser formulada de acuerdo con el principio de Mach depende de la hipótesis cosmológica que el investigador esté dispuesto a formular (junto, desde luego, con su capacidad matemática). Baste decir que la mayor parte de las cosmologías relativistas son incapaces de aceptar el principio de Mach enteramente, y el que pueda aceptarse *algo* de él es una difícil y polémica cuestión para los físicos matemáticos. Fijémonos brevemente en algunas de las dificultades con las que parece tenerse que enfrentar un cosmólogo que desee prescindir del espacio-tiempo absoluto [11].

Se recordará que, de acuerdo con la filosofía relacional del espacio expuesta por Mach, carece de importancia el que digamos que la Tierra gira con relación a las estrellas fijas, o que afirmemos que las estrellas fijas giran alrededor de la Tierra en órbitas determinadas por fuerzas apropiadas centradas en la Tierra. La teoría especial de la relatividad no tiene nada que decir sobre la cuestión, ya que trata únicamente de sistemas inertes, y el sistema de ejes con relación al cual la rotación de la Tierra es cero no cuenta como sistema inerte. Ahora bien, la teoría general de la relatividad se reduce, en efecto, a relatividad especial (dentro de los límites del error observacional) cuando tratamos de sistemas de

tamaño algo menor y siempre que estemos dispuestos a postular fuerzas adicionales si fuese necesario. Así, por poner un ejemplo ya muy gastado, considérese un ascensor acelerador. Podemos considerar al ascensor acelerador como a un ascensor no acelerador, es decir, como un sistema inerte, si introducimos una fuerza gravitacional adecuada. En lugar de considerar el suelo del ascensor empujando con una fuerza que nos acelera, nos consideramos empujados hacia el suelo mediante una fuerza gravitacional. Debe comprenderse, sin embargo, que tal reducción a la relatividad especial, que corresponde a un cambio de espacio-tiempo curvo a espacio-tiempo plano, puede hacer únicamente *localmente*. El caso es bastante semejante al de alisar una alfombra mal hecha. Podemos alisar las arrugas en una pequeña área, pero esto da lugar a arrugas en alguna otra parte. No podemos hacer que la alfombra quede lisa en todas partes [4].

Aceptar por igual las dos descripciones: *a*) ascensor acelerador y *b*) ascensor no acelerador, o *a'*) Tierra en rotación y *b'*) Tierra no en rotación, equivaldría a suponer que las arrugas pueden ser alisadas *en todas partes*. Si consideramos que la Tierra está en reposo nos encontramos con la dificultad de que la velocidad transversal de las estrellas tendría que ser mayor que la de la luz, lo cual resulta inconsistente con la teoría especial de la relatividad. No podemos, por tanto "alisar la alfombra", ni siquiera hasta la estrella más próxima. Grünbaum ha señalado [12], por consiguiente, que debe sostenerse que la Tierra gira no en relación con un sistema extensivo de ejes que llega hasta las estrellas fijas, sino con relación a un sistema local formado por los rayos de luz estelares conforme se aproximan a la Tierra, siendo determinadas sus rutas por el campo métrico local [5]. A menos que este campo métrico local pueda ser explicado de un modo puramente relacional, y no es, en modo alguno, fácil ver cómo podrá hacerse esto, no podemos tratar de la rotación de la Tierra con relación a las estrellas fijas de un modo que esté de acuerdo con el principio de Mach.

---

[4] Para evitar posibles malentendidos debe insistirse, desde luego, en que "curvo" y "plano", cuando se aplican al espacio-tiempo, reciben un significado matemático preciso que no depende de nuestra intuición, como sí depende nuestra analogía de la alfombra. La métrica en la relatividad general viene dada por:

$$\begin{aligned} ds^2 = {} & g_{11}\,dx^2 + g_{22}\,dy^2 + g_{33}\,dz^2 + g_{44}\,dt^2 \\ & + 2\,g_{12}\,dx\,dy + 2\,g_{13}\,dx\,dz + 2\,g_{14}\,dx\,dt \\ & + 2\,g_{23}\,dy\,dz + 2\,g_{24}\,dy\,dt + 2\,g_{34}\,dz\,dt \end{aligned}$$

Las *g*s son todas ellas funciones de (x, y, z, t). Decir que el espacio-tiempo es plano en un punto determinado es decir que los ejes x, y, z y t, pueden ser elegidos de tal modo que en un punto $g_{11}=g_{22}=g_{33}=-1$, $g_{44}=1$, y $g_{12}=g_{13}=g_{14}=g_{23}=g_{24}=g_{34}=0$. Obsérvese que el hecho de que en un sistema particular de ejes g_{12}, etc., *no* sean iguales a 0 en un punto determinado no demuestra que el espacio no es plano allí. De este modo el espacio bidimensional euclidiano es "plano" en todas partes debido a que pueden elegirse los ejes de tal modo que siempre $ds^2=dx^2+dy^2$, pero con ejes oblicuos tenemos $ds^2=dx^2+dy^2+kxy$. Un espacio es plano si existe *algún* sistema de ejes para el cual g_{12}, g_{13}, etc., sean cero. Véanse también págs. 184-187 posteriormente.

[5] Hablar del "campo métrico" es hablar de una determinada relación entre los g_{ik} (es hablar de las curvaturas en los diversos puntos del espacio-tiempo).

La razón por la que la teoría general de la relatividad esté tan ligada a la cosmología es la de que los presupuestos de tipo cosmológico relativos a las condiciones límite tienen que ser realizados antes de que puedan resolverse las ecuaciones de la teoría. Desde luego, los presupuestos relativos a las condiciones límite tienen que ser hechos en el caso de *cualquier* teoría, pero en la mecánica newtoniana, por ejemplo, éstos no tienen que ocuparse tanto del universo en su conjunto, ya que en la teoría newtoniana los cuerpos muy distantes tienen efectos imperceptibles entre sí. Un posible modo de solucionar las ecuaciones de Einstein es suponer que el espacio-tiempo es plano "en el infinito". Esto es algo así como alisar todos los bordes de la alfombra y concentrar las arrugas en el centro. Ahora bien, suponer que el espacio-tiempo tienen una estructura definida ("plana") infinitamente lejana a la materia no está, evidentemente, de acuerdo con el principio de Mach. Las ecuaciones modificadas de Einstein que proporcionan un universo finito pero no ligado no evitan la negación del principio de Mach tampoco, ya que implican una curvatura definida del espacio-tiempo incluso en ausencia de toda materia.

No está del todo claro, desde luego, que sea necesariamente deseable la reconciliación de la relatividad general con el principio de Mach. ¿Por qué no podría ser el espacio-tiempo una "entidad teórica" perfectamente válida por derecho propio, y no simplemente un sistema de relaciones entre otras entidades teóricas tales como electrones y protones? Hay una larga tradición en la física matemática que surge a partir de Descartes, y que debe ser contrapuesta a la filosofía relacional que surge a partir de Leibniz. La teoría del propio Descartes se fundaba en inconsistencias, ya que él era incapaz de hacer una distinción entre las partes del espacio, y no podía explicar cómo algunas partes del espacio consisten en cuerpos y cómo las otras partes del espacio pueden ser espacios *entre* cuerpos. Puesto que mantenía que la esencia de la materia es la extensión, tenía que negar la existencia de un vacío, pero su teoría no podía dar cuenta del hecho de que si no existe ningún vacío existe una diferencia entre las partes del espacio en que algunas, por lo menos, se acercan más a estar vacías que otras. Descartes también se vio obstaculizado por el hecho de que no había congregado al espacio y al tiempo juntos en una unidad espacio-temporal. Esto hizo que le resultara difícil el tratar del movimiento. A partir de los descubrimientos matemáticos de Riemann, sin embargo, la noción de un espacio de curvatura variable nos permite conceder un papel realmente físico al espacio, e incluso podemos jugar con la idea de que los cuerpos materiales no son sino singularidades en tal espacio-tiempo.

Otra idea de la matemática moderna de la que no dispuso Descartes es la del espacio múltiplemente conectado. Considérese la superficie de una esfera. Este es un espacio bidimensional. Considérese una curva cerrada sobre esta superficie. Tal curva puede estrecharse hacia abajo hasta convertirse en un punto. Supóngase ahora que nuestro espacio bidimensional es la superficie de un buñuelo. Es evidente que algunas curvas cerradas

en este espacio son tales que no pueden estrecharse hacia abajo hasta convertirse en un punto. La superficie de la esfera se dice que está "simplemente conectada". La superficie del buñuelo se dice que está "múltiplemente conectada". Una superficie está conectada n-múltiplemente si se necesitan $n-1$ cortes no-interseccionales para convertirla en simplemente conectada. Ahora, siguiendo esta analogía, trátese de concebir un espacio múltiplemente conectado de muchas dimensiones. No se le puede *imaginar,* pero confío que se le pueda *concebir.* Si se permiten tantas dimensiones como se deseen, cualquier grado de conexión, y curvatura variable, se tiene un tipo de espacio mucho más maravilloso que el espacio tridimensional euclidiano del que dispuso Descartes.

J. A. Wheeler [13] ha propuesto un modelo cosmológico semejante, según el cual los electrones y los protones son extremos opuestos de los "agujeros de gusano" en un espacio-tiempo múltiplemente conectado. (No es necesario en absoluto afirmar que existan estos agujeros de gusano. Sólo desde el punto de vista de un espacio tridimensional podemos afirmar que un buñuelo tiene un agujero en el medio. Desde el punto de vista de un ser bidimensional que exista en la superficie es una superficie bidimensional sin agujeros, pero con la propiedad, topológicamente interesante, de estar doblemente conectada.) La cosmología de Wheeler es, desde luego, especulativa, y así lo son probablemente todas las cosmologías, al menos hasta que llegue el día en que se dé una unificación de la microfísica (física de las partículas) con la teoría de campo de la relatividad general. En la actualidad es imposible decidir si la física llegará a explicar el espacio-tiempo según el modo de Leibniz, en términos de relaciones entre partículas o si será vindicada la idea cartesiana, y se considerará al universo construido, quizá según la manera ideada por Wheeler, por puntos espacio-temporales únicamente.

Geometría y física

Si el programa ideado por Wheeler fuese llevado a cabo finalmente, la teoría física terminaría siendo una geometría de curvatura variable de muchas dimensiones múltiplemente conectadas, fantásticamente complicada. Las complejidades de la geometría, sin embargo, quedarían compensadas por la extremada simplicidad ontológica de la teoría, ya que el mundo estaría construido por puntos espacio-temporales únicamente. Aunque la física no pase a ser totalmente geometría, si se necesita el espacio-tiempo absoluto en la teoría física, la geometría se convertirá en parte esencial de la teoría física, y los puntos espacio-temporales serán entidades teóricas a la par con cosas tales como electrones y protones. E incluso aun cuando la teoría relacional del espacio-tiempo resultase ser viable, el conjunto de relaciones entre las partículas físicas del que se trataría en geometría sería descriptible, únicamente, mediante leyes de la geometría que tendrían el status de leyes teóricas de la naturaleza.

La concepción acabada de mencionar está en contraste con la concepción operacionalista de la geometría descriptiva, que afirma que la geometría trata del comportamiento de traslación de varas rígidas. El que la geometría trate de varas rígidas puede ser refutado como sigue: El mundo podría haber sido puramente gaseoso, pero, con toda seguridad, aun en ese caso nosotros (quizá como personas gaseosas o sin cuerpo) podríamos haber utilizado la geometría para explicar las leyes de este mundo gaseoso. Puesto que no existirían varas rígidas, no utilizaríamos la geometría para explicar el comportamiento de traslación de varas rígidas. Por la misma razón, cuando la geometría *es* aplicada al comportamiento de varas rígidas, es mejor decir que la geometría *explica* tal comportamiento y no que lo *describe*. Es decir, la geometría tiene más contenido que la descripción de medidas con varas rígidas, al igual que la mecánica cuántica tiene más contenido del que podemos obtener sumando las informaciones de lo que el experimentador ve cuando mira los instrumentos del laboratorio.

Con todo, en cierto sentido podemos elegir nuestra geometría a nuestro gusto. Considérese, por ejemplo, algunas coordenadas cartesianas semieuclidianas (x, y, z, t). Es decir, podemos concebir nuestro diagrama de Minkowski cubierto con una malla cuadrada como en la figura 12. Ahora dibujemos cualquier malla que deseemos, como en la figura 13.

Hágase la transición de la figura 12 a la figura 13 de acuerdo con las siguentes ecuaciones, proporcionando $f, g, h, j,$ un continuo paralelismo biunívoco:

$$\begin{aligned} x' &= f(x, y, z, t) \\ y' &= g(x, y, z, t) \\ z' &= h(x, y, z, t) \\ t' &= j(x, y, z, t) \end{aligned}$$

También hay funciones $f', g', h', j',$ tales que

$$\begin{aligned} x &= f'(x', y', z', t') \\ y &= g'(x', y', z', t') \\ z &= h'(x', y', z', t') \\ t &= j'(x', y', z', t') \end{aligned}$$

Ahora si $Q(x, y, z, t)=0$ es una ley cualquiera de la naturaleza en las coordenadas originales, la sustituimos por.

$$Q\ \{f'(x', y', z', t'),\ g'(x', y', z', t'),\ h'(x', y', z', t'),\ j'(x', y', z', t')\}=0$$

De este modo es una cuestión sencilla cambiar conjuntamente nuestra geometría y nuestra física, de modo que la combinación de ambas continúe siendo esencialmente la misma teoría. Será consecuencia de la nueva

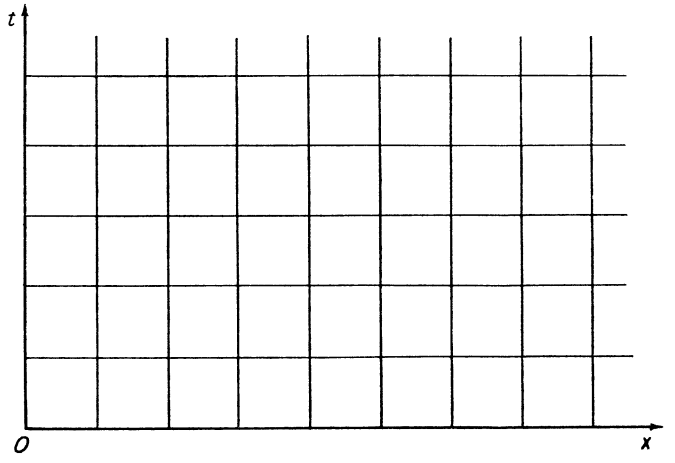

FIGURA 12.

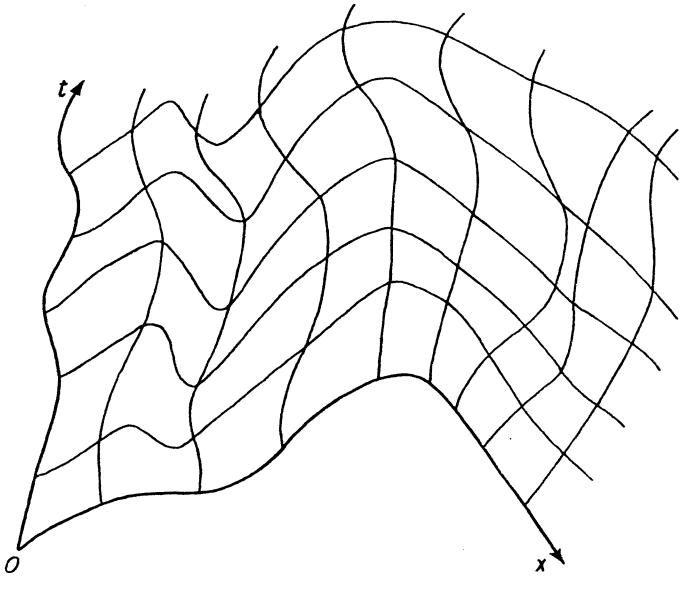

FIGURA 13.

teoría física y la nueva geometría que una vara que conserve en su traslación su longitud (cuando la longitud se mide en la geometría antigua) ya no conserve una longitud uniforme de acuerdo con la métrica de la nueva geometría. Podemos señalar esto del modo siguiente: las definiciones de congruencia de la nueva geometría son distintas de las de la antigua.

Por ejemplo, podemos modificar con facilidad nuestras leyes físicas de modo que pasemos de una geometría euclidiana (o más bien una geocronometría semi-euclidiana) a otra hiperbólica. Sin embargo, los cambios permisibles, en este sentido, en nuestra geometría están limitados por un requisito evidente. Es necesario conservar la topología de la geometría original. Así pues, podemos pasar de un espacio euclidiano a otro hiperbólico, debido a que éstos poseen la misma estructura topológica: son infinitos en todas las direcciones. No podemos pasar de un espacio euclidiano o hiperbólico a uno esférico o elíptico, o vice-versa, ya que un espacio esférico o elíptico es finito pero no ligado.

Supóngase que trasladamos un espacio esférico a otro euclidiano. Considérese una señal de luz que circunde el espacio esférico. Esta se sustituiría por una señal de luz que destellease en el espacio infinito seguida de una señal de luz que procediese del infinito a partir de la dirección opuesta. Admitir esto sería, como Reichenbach ha señalado [14] tolerar anomalías causales.

Por este motivo Grünbaum considera las propiedades topológicas como intrínsecas, mientras que considera a las propiedades métricas como relativas solamente a las convenciones de congruencia [15]. Tiene razón, sin duda, al considerar las propiedades topológicas como más fundamentales que las métricas, e incluso es posible que la física del futuro sea totalmente topológica y no métrica. Con todo, el punto de vista adoptado en los capítulos 3 y 5 quizá nos incline a considerar las definiciones de congruencia como no dadas *ab initio* sino como deducciones a partir de nuestra física y nuestra geometría. ¿Por qué no tomar esa combinación de física y geometría que lleva a la formulación más simple y suponer, basándonos en la mayor simplicidad, que ésta proporciona la imagen correcta del mundo? Quizá, sin embargo, esto sería conceder demasiada importancia al principio de simplicidad. Grünbaum diría que la simplicidad aquí es la simplicidad descriptiva solamente, no la simplicidad inductiva.

Supóngase que el espacio o el espacio-tiempo estuviesen constituidos por trozos discontinuos. Entonces no necesitaríamos las definiciones de congruencia (o teoremas de congruencia) de una geometría a fin de asignar una longitud a un objeto. La longitud de un objeto podría ser definida de modo muy natural como el número de trozos de espacio entre sus extremos. También podríamos considerar estos trozos como puntos y definir la longitud de cualquier cosa como cero, utilizando el criterio de la moderna teoría de medida, que asigna una medida cero a cualquier conjunto finito de puntos. Desde luego, si utilizásemos la teoría de la medida para nuestro criterio de longitud, cualquier objeto tendría longitud cero si el

espacio consistiese en un conjunto infinito de puntos meramente enumerables. Para espacios que (a consecuencia de cosas tales como la inconmensurabilidad de los lados y la hipotenusa de un triángulo isósceles con un ángulo recto) contienen una infinidad de puntos no enumerables, la teoría de la medida asigna una longitud no cero a un intervalo. Las longitudes reales que tienen que ser asignadas ahora dependen de reglas de congruencia o teoremas de congruencia.

Desde un punto de vista operacional definimos "rígido" en términos de "no rígido". Así, evidentemente, un trozo de caucho no es rígido porque aunque sus extremos coincidan en un momento dado con los extremos de un trozo de hierro o madera, también pueden no coincidir, incluso aunque el caucho esté tirante. Sin embargo, el hierro y la madera son aproximadamente rígidos, porque si los extremos coinciden una vez entonces coinciden siempre. En un grado mayor de aproximación puede comprobarse que el hierro y la madera son ellos mismos "no rígidos", ya que la madera puere dejar de coincidir con el hierro después de haber sido humedecida. Igualmente, piezas de hierro y bronce cuyos extremos coinciden a una temperatura no coinciden a otras temperaturas. Puede darse una definición más matizada en términos de hierro y bronce a una determinada temperatura (por ejemplo el punto de congelación del agua). Si los trozos de hierro y bronce (que estén a esta temperatura) coinciden en un lugar o tiempo coincidirán en todos los lugares y tiempos. Más tarde podremos utilizar nuestras burdas medidas de longitud para construir instrumentos de laboratorio, tales como espectómetros, y entonces podemos utilizar una medida de longitud todavía más matizada, tal como la de la longitud de onda de la luz de cesium. Ahora bien, podemos establecer reglas de correspondencia entre nuestra geometría relativista y nuestras medidas operacionales de longitud, y decir que nuestra medida operacional y nuestra medida geométrica coinciden en regiones donde no hay un campo gravitacional demasiado fuerte.

Grünbaum considera las reglas de correspondencia rudimentarias de las que acabo de hablar como *definiciones de congruencia.* Mantiene entonces que la geometría es una ciencia empírica en el sentido de que una geometría viene a ser especificada únicamente por los hechos del mundo juntamente con nuestras definiciones de congruencia. Si así lo deseamos, podemos estipular que la longitud de una vara rígida sea alguna función $f(x, y, z, t)$ de su posición espacio-temporal, en cuyo caso pasamos a una geometría diferente, y que las g_{ik} de la tórmuli $ds^2=g_{ik}dx_i dx_k$ sean diferentes. Con todo, Grünbaum mantendría que ésta no es una geometría distinta en el sentido de que su significado físico sea distinto del de la antigua. El cambio en las definiciones de congruencia y el de las g_{ik} se compensan entre sí. De este modo Grünbaum utiliza el caso de la geometría física como contra-ejemplo de la tesis duhemiana de que se puede mantener una teoría a la vista de testimonios en contra, con tal que verifiquemos cambios adecuados en nuestros presupuestos básicos. La respuesta de Grünbaum [16] es que una geometría *física* (es decir, la teoría matemá-

tica *junto con* las definiciones de congruencia) está determinada únicamente por los hechos empíricos, al menos si podemos ignorar los habituales riesgos inductivos de extrapolar de lo que observamos en un determinado número de observaciones a lo que observaríamos si realizásemos observaciones posteriores indefinidamente. (Esta última puntualización carece de importancia, ya que es defícil ser duhemiano acerca de "todos los cuervos son negros", cuando también se da riesgo inductivo.)

Parece dudoso, sin embargo, que tal concepción inductivista de la geometría pueda valer cuando la geometría se considera no como una descripción del comportamiento de traslación de cuerpos rígidos, como se le considera vulgarmente, sino *explicación* de esto. Considérese, por ejemplo, la regla de congruencia: "La vara métrica de París tiene la misma longitud en todas partes." Esto podría ser considerado como una proposición analítica, es decir, aquélla que es verdadera en virtud de convenciones lingüísticas únicamente. Pero ahora que contamos con una regla de longitud basada en la longitud de onda de la luz de cesium, la proposición relativa a la vara métrica de París resulta evidentemente fáctica y deja de ser ya meramente convencional. Incluso desechando esta consideración, es físicamente significativo que la vara métrica de París pueda doblar su longitud: como Hilary Putnam [17] ha señalado, desde el punto de vista de la mecánica cuántica es del todo posible el que una vara rígida doble espontáneamente su longitud (aunque, desde luego, infinitamente improbable, de modo que no estaría jamás justificado que creyésemos una información relativa a tal alegado acontecimiento). Incluso la longitud de onda de la luz de cesium es una regla de longitud solamente si no nos encontramos en un campo gravitacional demasiado intenso. Puede verse, por consiguiente, que una definición puramente operacional de la longitud es extraña a la ciencia moderna, y que lo que realmente ocurre es que la medida de longitud (el tensor métrico espacio-temporal) se postula dentro de una teoría, y la teoría está ligeramente ligada a la experiencia mediante reglas de correspondencia aproximadas (e incluso éstas serían innecesarias en el caso de una extinción feyerabendiana de la ciencia, en la cual nuestro discurso observacional sería realizado todo él dentro del propio lenguaje teórico).

Una razón por la que Grünbaum parece negar que la geometría sea una teoría explicatoria análoga a la teoría electromagnética, por ejemplo, es la de que con una transformación adecuada de la forma matemática de las leyes físicas podemos contar (con tal que elijamos una geometría con la topología adecuada) con cualquier tipo de geometría que gustemos. (Véase la sustitución de la red rectangular por la ondulada en las figuras 12 y 13, anteriormente.) Pero, replica Putnam [18], ¿no podríamos, del mismo modo, volver a medir algunas cantidades admitidas como *físicas,* por ejemplo sustituir la presión P por una nueva medida $f(P)$? Donde P aparecía en una ley física $f'(P)$ aparecerá ahora donde f' sea la función inversa de f. ¿Significa esto que la presión también es "métricamente amorfa"? De no serlo, entonces un intervalo espacio-temporal puede ser

considerado una cantidad física exactamente igual que puede serlo la presión.

Me da la impresión de que nos estamos metiendo en un problema que podría ser evitado tan sólo con que *rechazásemos* la distinción de Grünbaum entre la simplicidad meramente descriptiva y lo que él denomina simplicidad "inductiva". Considérese, en particular, la ley: "la línea del mundo de un cuerpo no accionado por ninguna fuerza (no gravitacional), por ejemplo, un rayo de luz o un satélite espacial, es geodésica" [6]. En una nueva asignación de reglas métricas del espacio-tiempo, hecha arbitrariamente, esto tendría que ser sustituido por algo mucho más complicado, y, con toda seguridad, perdería su fuerza explicatoria. Ya que dejaría de aparecer como un principio máximo o mínimo (análogo, por ejemplo, a la primera ley del movimiento de Newton, en la física clásica), sino que aparecería como algo confuso e increíble, salvo en el supuesto de que hayamos asignado medidas del espacio-tiempo de tal modo que resulte más *oscuro* el modo en que el espacio-tiempo existe en realidad. En cuyo caso ¿no podemos decir que el espacio-tiempo *es en realidad* tal que para una partícula libre $\delta \int ds = 0$ (es decir, que su línea del mundo es geodésica)?

Nuestra concepción de la geometría es la de una hipótesis física ligada un tanto vagamente a la experiencia mediante reglas de correspondencia que *no* son definiciones y *no* son inviolables. En este sentido no se diferencia de la teoría, por ejemplo, de que la materia está compuesta de determinados tipos de partículas subatómicas. Hay que añadir unas cuantas observaciones respecto a la geometría, tal como se la concibe desde este punto de vista, relacionada con la geometría tal como es considerada por los matemáticos puros que estudian tipos de espacios, sin preocuparse de si se realizan físicamente o no.

Es necesario que distingamos, por consiguiente, entre una geometría *a priori* (la geometría tal como se presenta al matemático puro) y una geometría física. Hasta la obra de W. Bolyai y N. I. Lobachevski en la primera mitad del siglo diecinueve no fue comprendida esta distinción entre la geometría pura y la geometría física, ya que se creía que existía solamente una geometría, la geometría euclidiana, y, por consiguiente, un estudio puramente matemático de la geometría no dejaba de ser aplicable al espacio real. Sin embargo, Bolyai y Lobachevski mostraron, independientemente, que era posible desarrollar una geometría en la cual el axioma de las paralelas de la geometría euclidiana es sustituido por el axioma de que a través de cualquier punto P que no esté en la línea L, existe más de una paralela a L. (Existen dos líneas a través de P que son denominadas paralelas a L, y, además, cualquier línea de la infinita cantidad de ellas que puede haber entre estas paralelas no se entrecruza con L.) Dándole

[6] Una geodésica es o bien la distancia más corta o la más larga entre dos puntos. Así pues, en la superficie de la Tierra las geodésicas son grandes círculos. Debido a que g_{11}, g_{22} y g_{33} son geodésicas negativas en el espacio-tiempo de la relatividad, son las distancias más largas.

una interpretación dentro de la geometría euclidiana, es posible demostrar qu la geometría de Bolyai y Lobachevski (geometría "hiperbólica") es consistente si la geometría euclidiana lo es. Palabras tales como "línea recta" tienen, desde luego, un significado distinto dentro de aquella interpretación, pero cuando les damos a estas palabras los nuevos significados euclidianos vemos que los axiomas de la geometría hiperbólica siguen siendo verdaderos.

Otro modo de negar el axioma de las paralelas es decir que no existen *ningunas* paralelas respecto a una línea determinada a través de un punto que no esté en ella. Si la geometría está completamente axiomatizada, como en el caso de la axiomatizada por David Hilbert, es ahora necesario también un cambio consiguiente en algunos de los axiomas puramente topológicos. Así pues, no será verdadero que si *B* está entre *A* y *C*, entonces *C* no esté entre *A* y *B*. De este modo tenemos la geometría elíptica.

Las geometrías hiperbólicas y elípticas son, al igual que la geometría euclidiana, geometrías de curvatura constante. Desde luego, estamos hablando acerca de espacios tridimensionales ahora, y la curvatura de tales espacios no puede ser visualizada intuitivamente, al igual que podemos visualizar el aspecto curvo de una esfera o un hiperboloide. A la curvatura del espacio en un punto *P* se le da una definición puramente matemática. Podemos ver cómo opera si consideramos un caso bidimensional. Supóngase que dibujamos un triángulo con un ángulo recto (cuyos lados sean grandes círculos) sobre la superficie de una esfera. Se descubrirá que el teorema de Pitágoras no es obedecido. Así pues, la "curvatura" de la superficie de la esfera podría ser descubierta por los seres que viviesen únicamente sobre la superficie.

En el caso euclidiano tenemos $ds^2=dx^2+dy^2$. donde ds es la longitud de la hipotenusa de un pequeño triángulo cuyos lados tienen la longitud dx y dy. En el caso más general tenemos $ds^2=g_{11}dx^2+2g_{12}dx\,dy+g_{22}dy^2$. Si es un espacio de "curvatura variable" g_{11}, g_{12} y g_{21} serán funciones de x e y. La "curvatura de un punto" puede ser definida simplemente en términos de g_{11}, g_{12} y g_{22}, y es igual a la curvatura gaussiasa $\frac{1}{R_1R_2}$, donde R_1 y R_2 son los principales radios de curvatura de un punto. En el caso bidimensional los propios R_1 y R_2 pueden ser visualizados, pero aunque puede hacerse esto en el caso de tres o más dimensiones, podemos seguir definiendo la curvatura en términos de g_{11}, g_{12}, etc. Así pues, en el caso cuatridimensional es posible definir la curvatura en términos de los g_{11}, g_{12}, etc., que aparecen en la fórmula:

$$\begin{aligned} ds^2=&g_{11}dx^2+g_{22}dy^2+g_{33}dz^2+g_{44}dt^2\\ &+2g_{12}dxdy+2g_{13}dxdz+2g_{14}dxdt\\ &+2g_{23}dydz+2g_{24}dydt+2g_{34}dzdt \end{aligned}$$

Cuando las *g*s son funciones de x, y, z, t, nos encontramos en el caso de

un espacio de curvatura variable, necesitado por la teoría general de la relatividad.

Hemos observado que puede demostrarse que la geometría hiperbólica es consistente, si la geometría euclidiana lo es, interpretándola dentro de la geometría euclidiana. Sin embargo, si se supone que el análisis, es decir, la teoría matemática del número real, es consistente, es una cuestión sencilla demostrar la consistencia de cualquier geometría, euclidiana o no euclidiana, de curvatura constante o de curvatura variable. Los puntos se definen simplemente (por ejemplo, en el caso tri-dimensional) como triples de números reales, los planos como clases de triples de números reales que satisfacen las ecuaciones de primer grado, las líneas como clases de triples de números reales que satisfacen pares de ecuaciones de primer grado, y así sucesivamente. De acuerdo con tal re-interpretación puede verse que los axiomas de la geometría son verdaderos, y, por tanto, la geometría deberá ser consistente si el análisis lo es.

En el caso de una axiomatización de geometría pura, tal como la de Hilbert, se dice comúnmente que los puntos, líneas y planos son "cualquier cosa que satisfaga los axiomas". Con más exactitud, Nagel [19] ha caracterizado los axiomas de la geometría pura como "formas de enunciado", es decir, en la terminología del capítulo 2, anteriormente, esquemas de sentencias. Así, de acuerdo con el punto de vista de Nagel, el geómetra puro muestra simplemente que determinados esquemas de sentencias pueden ser deducidos de otros esquemas determinados. Esto podría ser demostrado expresando los axiomas en la terminología del cálculo de predicados y utilizando únicamente determinadas letras de sentencias específicas en lugar de "punto", "línea", "plano", etc. Se vería entonces que en tanto que a estas letras de sentencias no se les diese una interpretación, el geómetra puro no haría ninguna afirmación y no tendría materia de qué tratar.

Sin embargo, si el geómetra puro no desea considerarse sin materia de qué tratar, siempre puede proporcionarse un tema que es independiente de la naturaleza del espacio físico, interpretando sus esquemas dentro de la teoría de los números reales. De este modo "es un punto" dentro de una geometría tridimensional dejará de ser un mero esquema (como "*Fx*") y se convertirá en el predicado concreto "es un triple de números reales". La geometría que se desee en particular se obtiene especificando las funciones que corresponden a las *g*s en la ecuación $ds^2=\Sigma g_{ik}dx_i dx_k$.

En una física geometrizada, tal como vimos ideó J. A. Wheeler, los puntos espacio-temporales son postulados como entidades teóricas: se convierten en la materia última del mundo. Zenón formuló una vez la cuestión de cómo los puntos no dimensionales podían dar lugar a figuras extensas. Grünbaum ha aplicado la moderna teoría de la dimensión para resolver esta paradoja [20]. De acuerdo con la teoría de la dimensión cualquier conjunto de puntos finito o enumerable tiene la dimensión cero, pero un conjunto de puntos *no-enumerable* (continuo) puede tener la di-

mensión uno o mayor que uno. (Hablando estrictamente, es a conjuntos de puntos, no a los puntos, a lo que abscribimos dimensiones, y, por tanto, el término "puntos sin dimensión" carece en realidad de significado. Lo que se quiere decir al hablar de puntos sin dimensiones quizá podría expresarse mejor hablando de unidades de conjuntos de puntos de dimensión cero.)

Por tanto, la cuestión de "¿de qué trata la geometría?" puede ser contestada de diversos modos. Tal como se da en la física geometrizada, trata de puntos espacio-temporales. En una física de partículas que haga uso enteramente de un concepto relacional del espacio y el tiempo, la geometría se incorpora en una física que trata de partículas únicamente, y las posiciones de relación aparecen como relaciones entre estas partículas [7]. En el caso de la geometría axiomática pura, los axiomas y los teoremas pueden ser considerados simplemente como esquemas de sentencias, pero si se desea considerarlos como enunciados, puede dárseles una interpretación dentro de la teoría de los números reales. La geometría analítica es o bien geometría física (acerca de puntos espacio-temporales postulados) o es geometría pura, en cuyo caso trata acerca de conjuntos de números reales.

Sugerencias respecto a lecturas posteriores

Una buena relación de la historia de las teorías filosóficas y científicas del espacio puede verse en la obra de Max Jammer, *Concepts of Space*, edición revisada (New York: Harper and Row, 1960). *The Philosophy of Space and Time*, de Hans Reichenbach (New York: Dover, 1958), contiene una útil discusión de las geometrías no euclidianas y del espacio-tiempo en la relatividad especial y general. Discusiones más recientes son las de Adolf Grünbaum, *Philosophical Problems of Space and Time* (New York: Alfred A. Knopf, 1963) y O. Costa de Beauregard, *La Notion de Temps, équivalence avec l'éspace* (París: Hermann, 1963). Una obra incluso más reciente de Grünbaum, que no pudo ser discutida en el presente volumen, es su *Modern Science and Zeno's Paradoxes* (Middleton, Conn.: Wesleyan University Press, 1967). Divulgaciones un tanto antiguas, pero todavía útiles, de la teoría de la relatividad, pueden verse en E. Borel, *Space and Time* (New York: Dover, 1960), publicada originalmente en una edición francesa, 1922, y Sir Arthur Eddington, *Space, Time and Gravitation* (New York: Harper and Row, 1959), publicada originalmente por Cambridge University Press, en 1920. El primer trabajo de Einstein sobre la relatividad "On the Electrodynamics of Moving Bodies" y el trabajo de Minkowski "Space and Time" hacían uso de matemáticas muy simples únicamente, y pueden verse en *The Principle of Relativity, A Collection of Original Memoirs on the Special and General Theory of Relativity*, por H. A. Lorentz, A. Einstein, H. Minkowski y H. Weyl, con notas de A. Sommerfeld, y traducida por W. Perrett y G. B. Jeffrey (London: Methuen, 1923, y en rústica, New York: Dover Books, 1952). Una colección de ensayos sobre la filosofía del espacio y el tiempo, editada con una introducción a cargo de J. J. C. Smart en *Problems of Space and Time* (New York: Macmillan, 1964).

[7] Para ver cómo puede hacerse esto, consúltese W. V. Quine, *Word and Object*, sección 52 [21].

Esta obra contiene una re-edición del trabajo de Minkowski y también selecciones de Leibniz, Newton y Mach acerca de las teorías absolutas y relativas del espacio. También contiene un útil trabajo de H. P. Robertson, "Geometry as a Branch of Physics"; parte del ensayo de Reichenbach, "The Philosophical Significance of the Theory of Relativity"; y un extracto de las notas bibliográficas de Einstein, todas las cuales aparecieron originalmente en P. A. Schilpp (ed.), *Albert Einstein, Philosopher-Scientist* (La Salle, Ill.: Open Court, 1951, y en rústica, New York: Harper and Row, 1959). Las teorías del espacio y el tiempo de Newton fueron criticadas por George Berkeley en el siglo dieciocho en sus *Principles of Philosophy,* secciones 110-116, y su *De Motu.* Un estudio de Berkeley como precursor de Mach y de Einstein aparece en el capítulo 6 de la obra de K. R. Popper, *Conjectures and Refutations* (London: Routledge and Kegan Paul, 1963, y New York: Basic Books, 1963). El estudio de Leibniz de la teoría relacional del espacio y el tiempo puede verse mejor que en ninguna otra parte en *The Leibniz-Clarke Correspondence,* que ha sido editada por H. G. Alexander (Manchester: Manchester University Press, 1956, y New York: Barnes and Noble, 1964). Clarke fue un defensor de Newton. La introducción del propio Alexander a este volumen es de considerable interés.

El capítulo 9 de la obra de Ernest Nagel, *Structure of Science* (New York: Harcourt Brace and World, 1961), contiene una discusión de geometrías alternativas, con un esbozo del intento de un tratamiento unificado de las geometrías no euclidianas desde el punto de vista de la geometría proyectiva. Sobre esta cuestión, véase Felix Klein, *Elementary Mathematics from and Advanced Standpoint,* traducida de la tercera edición alemana por E. R. Hedrick y C. A. Noble (New York: Dover, 1939), parte 3, especialmente págs. 174-184.

Sobre la paradoja del reloj hay una nota muy concisa y acertada de W. H. McCrea en *Discovery,* XVIII (febrero, 1957), págs. 56-58, y otra acertada discusión es la de las págs. 215-222 y 231-232 de la obra de G. J. Whitrow, *Natural Philosophy of Time* (London: Nelson, 1961, y en rústica, New York: Harper and Row, 1963). Una discusión no geométrica es dada por Adolf Grünbaum en su trabajo "The Clock Paradox in the Special Theory of Relativity", *Philosophy of Science,* XXI (1954), 249-253, y las respuestas a discusiones de esto, *Philosophy of Science,* XXII (1955), 53, 233. Sobre la paradoja métrica de la extensión de Zenón (la cuestión de cómo construir un espacio y un tiempo extensos a partir de elementos no extensos), véase el capítulo 6 de la obra de Grünbaum, *Philosophical Problems of Space and Time, op. cit.,* o su tratamiento más detallado en "A Consistent Conception of the Extended Linear Continuum as an Aggregate of Unextended Elements", *Philosophy of Science,* XIX (1952), 288-306 y en el capítulo 3 de su libro *Modern Science and Zeno's Paradoxes* (Middletown, Conn.: Wesleyan University Press, 1967). El trabajo de Hilary Putnam, "An examination of Grünbaum's Philosophy of Geometry" en B. Baumrin (ed.), *Philosophy of Science, The Delaware Seminar,* Vol. 2, 1962-63 (New York: Interscience, 1963), páginas 204-255, no es siempre del todo justo con Grünbaum, como cuando critica la discusión de Grünbaum de la paradoja métrica de la extensión de Zenón como si estuviera hablando acerca de las paradojas, completamente distintas, del movimiento de Zenón, pero, como discusión independiente de la filosofía del espacio es muy valiosa. La respuesta de Grünbaum al trabajo de Putnam puede verse en el volumen de próxima aparición de *Boston Studies in the Philosophy of Science,* y en el libro de Grübaum de próxima aparición, *Geometry and Chronometry in Philosophical Perspective* (Minneapolis: University of Minnesota Press).

Un excelente trabajo de divulgación de preguntas sobre la simetría en física (incluyendo la caída de la paridad) puede verse en el ingenioso libro de Martin Gardner, *The Ambidextrous Universe* (New York: Basic Books, 1964). Véase también Eugene P. Wigner, "Violations of Symmetry in Physics", *Scientific American,* CCXIII (diciembre, 1965), 28-36.

8. LA DIRECCION DEL TIEMPO

El supuesto paso del tiempo

En el último capítulo adoptamos la posición de que el universo está constituido por un espacio-tiempo cuatridimensional. El presente, el pasado y el futuro son todos ellos igualmente reales. Cuando decimos que un acontecimiento está en el futuro estamos diciendo que es posterior a nuestras palabras, cuando decimos que es presente estamos diciendo que es simultáneo con nuestras palabras y cuando decimos que es pasado estamos diciendo que es anterior a nuestras palabras. El presente, por supuesto, participa de la relatividad de la simultaneidad, a menos que, como parece plausible, interpretemos el presente como la simultaneidad con una expresión lingüística relativa al sistema en el que quien la pronuncia está en reposo. Sin embargo, estas sutilezas no interesan demasiado. "Presente", "pasado" y "futuro" son términos de nuestro lenguaje cotidiano, y a la ciencia puede irle mucho mejor sin ellos. Una objeción relativa a las sentencias "E es presente", "E es pasado" y "E es futuro" es que son dependientes de un contexto: sus valores de verdad depende del momento en que son pronunciadas. Para fines científicos "E es simultáneo con E_0", siendo E_0 algún acontecimiento de referencia, va mucho mejor. O mejor todavía "E está en t", siendo t la coordenada temporal de E. Dentro de la teoría científica podemos utilizar "está" sin referencia a tiempo verbal: "el eclipse de la Luna está en t", cuando se dice antes o después de t, no significa, desde luego, que el eclipse de la Luna esté en *ese momento* en t.

Ciertamente tenemos la impresión de que el concepto del presente es metafísicamente importante: parece ser la parte superior de la ola en la que (al igual que los que practican el esquí acuático) somos empujados hacia el futuro. Tales impresiones metafísicas son muy confusas: es absurdo afirmar que nos estamos adentrando en el futuro; no hay ola que pueda correr hacia adelante en el espacio tiempo, puesto que el tiempo está ya en el espacio-tiempo. Hablamos del movimiento a través del espacio y, desde el punto de vista del espacio-tiempo se trata únicamente de inclinaciones relativas de las líneas del mundo, ya que mi movimiento con relación a ti es exactamente la proporción de cambio de la distancia entre nosotros respecto al tiempo. Evidentemente, no hay lugar en la imagen del espacio-tiempo para movimiento *a través* del espacio-tiempo [1].

Puesto que el movimiento es el cambio de espacio con respecto al tiempo, ¿qué podría ser el movimiento a través del tiempo? ¿Cambio de

---

[1] Una exposición clara de este punto puede verse en Moritz Schlick, *Philosophy of Nature,* capítulo 7 [1].

tiempo con respecto a qué? Nuestra impresión de que nos estamos moviendo a través del tiempo [2] es, por consiguiente, una ilusión metafísica basada en una confusión mental. A primera vista podría ser explicada simplemente por el hecho de que nuestro caudal de recuerdos está aumentando siempre, de que, por ejemplo, dentro de cinco minutos tendremos más recuerdos de los que tenemos ahora. Si nos identificamos con nuestro caudal de recuerdos quizá tengamos la impresión de ser prolongados continuamente en la dirección del tiempo. Esto, sin embargo, no es válido. Si tenemos más recuerdos dentro de cinco minutos de los que tenemos ahora, ¿por qué no experimentamos la sensación de hundirnos en el pasado con la misma facilidad con que nos sentimos prolongados hacia el futuro? O, también, considérese el caso de un hombre muy anciano, cuyos recuerdos se van extinguiendo. ¿No estará realmente *disminuyendo* su caudal de recuerdos, puesto que la pérdida de los mismos sobrepasa a la adquisición de otros nuevos? La explicación de la confusa impresión que experimentamos de avanzar a través del tiempo no puede ser simplemente el que tengamos *más* recuerdos en t_2 que en t_1 (siendo t_2 posterior a t_1). Quizá la siguiente sugerencia pueda proporcionar la respuesta: En t_2 tenemos recuerdos de diversos acontecimientos A_1, A_2, A_3, ..., y algunos de estos acontecimientos son ellos mismos recuerdos de otros acontecimientos del conjunto. Supóngase, ahora, que A_1, A_2, A_3, ..., están relacionados mediante una relación triádica relativa a cuál está entre cuál, de tal modo que si A_k está entre A_h y A_l, A_h no está entre A_k y A_l, y A_l no está entre A_h y A_k. De este modo podemos colocar a A_1, A_2, A_3, ..., en *orden,* pero esto no contiene un sentido intrínseco de la dirección. Tomemos, sin embargo, arbitrariamente, una dirección tal que la dirección de A_n a A_{n+1} sea la dirección positiva. Quizá nos encontremos con que hemos elegido la dirección de tal modo que si A_q es un recuerdo de A_p, la dirección de A_p a A_q es siempre positiva. Entonces nuestra dirección positiva definida arbitrariamente resultará ser la misma que la de lo anterior a lo posterior. Ahora bien, el mero hecho de que seamos conscientes de lo anterior y lo posterior no explica nuestra sensación subjetiva de progreso a través del tiempo. ¡También podemos distinguir el norte del sur sin tener la impresión de que algún día seremos esquimales! Grünbaum, tal como yo lo entiendo, ha sugerido que nuestra impresión de progreso a través del tiempo se deriva del sentido de dirección impuesto por el hecho de que nuestros recuerdos incluyen recuerdos de recuerdos, y esto impone un sentido de dirección (del acontecimiento recordado al recuerdo de él, al recuerdo de recordarlo) que presenta una vidente asimetría [3].

---

[2] O de que acontecimientos futuros se acercan a nosotros.

[3] Baso las observaciones precedentes en algunas sugerencias de la pág. 325 de la obra de Adolf Grünbaum, *Philosophical Problems of Space and Time* [2], y en el enunciado ligeramente más completo de la pág. 663 del trabajo de Grünbaum "Carnap's views on the Foundations of Geometry" [3], donde Grünbaum dice: "El flujo del tiempo consiste en la *conciencia instantánea tanto* del orden temporal *como* de la *diversidad* de miembros pertenecientes al conjunto de acontecimien-

Sean o no correctas las ideas acabadas de mencionar, parece ser, en cualquier caso, que nuestra idea errónea del flujo del tiempo tiene algo que ver con la asimetría inherente a la memoria: que los recuerdos se refieren a acontecimientos que ocurrieron antes que los recuerdos, no con posterioridad a ellos. El que los recuerdos se refieran a acontecimientos anteriores es un caso especial de un fenómeno muy general, que es el de la asimetría de los registros y las huellas. De este modo, hoy un magnetófono puede tener en su cinta una gravación de la reunión filosófica de ayer, pero no puede tener una gravación de la reunión de mañana. Las huellas de pisadas de hoy, dejadas entre los rosales, pueden ser la señal del robo de la semana pasada, pero no de la que viene. Un fósil en unas rocas puede constituir la huella de un animal que vivió hace un millón de años, pero no hay nada en las rocas que nos informe acerca de futuras especies. Los propios recuerdos consisten en determinadas huellas físicas en el cerebro [4]. ¿Podríamos encontrar lo que es común a todos los casos de registros o huellas, ya se trate de huellas cerebrales, gravaciones en cinta magnetofónica, huellas de pisadas, fósiles, o cualquier otra cosa, y que explique por qué proporcionan un conocimiento más o menos válido de acontecimientos anteriores, no de posteriores? La asimetría epistemológica parece ser un símbolo de la ampliamente difundida asimetría física del universo, o al menos, de nuestra parte de él.

Es cierto, desde luego, que conocemos bastante con relación al futuro. Un astrónomo puede predecir las posiciones de los planetas, y, de un modo bastante menos seguro, podemos predecir tormentas a partir de la información metereológica de la prensa, y a partir del crujido de una viga podemos predecir la caída de un puente. Estos casos de predicción son aquellos que pueden ser subsumidos por el modelo legal de cobertura" (véase capítulo 3), y en el caso de predicciones a partir de leyes se trata, indudablemente, de una simetría entre el pasado y el futuro. El modo en que un astrónomo predice la posición de Júpiter dentro de seis meses no se diferencia del modo en que determina la posición de Júpiter hace seis meses. En resumen, es evidente que nuestro conocimiento del pasado mediante huellas es bastante diferente de la retrodicción mediante leyes, pero aunque no se esté de acuerdo con esto debe admitirse, con toda seguridad, que podemos poseer un conocimiento mucho más extenso del pasado de un modo en que no podemos tener conocimientos detallados del futuro, y que debe existir alguna asimetría en el universo, o al menos en nuestra parte de él, que explique este hecho conocido.

---

tos recordados (registrados) u olvidados, conciencia en cada uno de los cuales el instante de su propio acontecer constituye un *elemento diferenciable*."

[4] Escribo como materialista. Incluso el dualista psico-físico admitiría probablemente que si los recuerdos no son huellas físicas dependen de ellas, y por tanto la disputa del materialismo en contra de otras teorías de mente-cuerpo no es importante aquí.

La asimetría temporal y la mecánica estadística

¿De dónde procede esta asimetría temporal? Podría sugerirse que la asimetría proviene de las leyes físicas. Esto parece poco probable. Las leyes de la mecánica y la electrodinámica clásicas son simétricas temporalmente: es decir, alguna función de $+t$ es una solución a ellas, también lo es la misma función de $-t$. (Esto se debe a que sus derivadas temporales son todas ellas de segundo orden.) Esta característica de la mecánica clásica se sigue conservando en la mecánica cuántica ordinaria. (Es cierto que la ecuación de Schrödinger contiene una derivada de primer orden, pero ésta es multiplicada por la raíz cuadrada de menos uno. Si eliminásemos la raíz cuadrada de menos uno, tendríamos que elevar al cuadrado la derivada de primer orden, y ello muestra que se sigue conservando la simetría temporal en la ecuación de Schrödinger.) La termodinámica fenomenológica no es, desde luego, temporalmente simétrica, pero hemos desechado la idea de que ésta sea un teoría fundamentalmente correcta. Considérese, por ejemplo, la "ley" fenomenológica de que un cuerpo que se deslice por una superficie dura perderá velocidad debido a la fricción. No ocurre nunca en nuestra experiencia que se dé el fenómeno contrario, a saber, que un cuerpo en reposo sobre una superficie dura empiece a moverse y aumente gradualmente su velocidad. Sin embargo, esto no se debe a que sea contrario a las leyes de la naturaleza, sino a que es muy poco probable. Imagínense a las moléculas de la superficie dura saliendo a golpear a las moléculas del objeto deslizante, y a las moléculas del objeto deslizante saliendo a golpear a las moléculas de la superficie dura. Inviértanse todas estas velocidades y se obtendrá el caso del objeto en reposo comenzando a moverse, con un enfriamiento de la superficie y del objeto.

Habría que hacer una matización respecto a la afirmación de que las leyes fundamentales de la naturaleza son temporalmente simétricas. Un experimento de J. H. Christenson, J. W. Cronin, V. L. Fitch y R. Turlay [4], del que se informó en 1964, sugiere, en realidad, una posible violación de la simetría temporal (aunque sigue habiendo simetría si invertimos no solamente el tiempo sino también la carga y la paridad). Sin embargo, me da la impresión de que no es del todo cierto que el único modo válido de explicar el resultado de este experimento sea la postulación de una violación de la simetría temporal. Más aún, es difícil ver cómo una violación de la simetría temporal tan recóndita, pueda tener mucho que ver con la muy ostensible asimetría temporal de los registros o de las huellas, de los que nos ocupamos en este capítulo.

Considérese una baraja que sea dispuesta inicialmente en grupos. Supongamos que las cartas son indestructibles y que son barajadas por una persona inmortal. Si se continúa barajando el tiempo suficiente (un tiempo extraordinariamente largo, por supuesto) la baraja volverá otra vez a su estado inicial, estando ordenada en grupos. Desde luego, el barajar las

cartas hará que el estado inicial vuelva una y otra vez, y finalmente incluso podrá darse el caso de que las cartas vuelvan a su estado ordenado a partir de un estado muy desordenado mediante, precisamente, la secuencia inversa de aquellos pasos intermedios que llevaron a las cartas de su estado ordenado a su estado de bien barajadas. En el tiempo infinito un acto particular de barajar no es más ni menos probable que la imagen exacta reflejada en el tiempo de este acto de barajar en particular. La razón por la que consideramos al barajar como un proceso irreversible se debe simplemente a que si partimos de una disposición en grupos y barajamos durante un tiempo especificado es casi seguro que las cartas acabarán por estar "bien barajadas". Y esto se debe a que el número de estados que estamos dispuestos a denominar como de cartas bien "barajadas" es incomparablemente mayor al número de estados que estamos dispuestos a denominar de cartas "ordenadas en grupos".

Considérese ahora un recipiente que contenga dos gases en compartimentos separados, divididos por una pared. Si sacamos la pared los gases se mezclarán. Podemos estar completamente seguros de ello, ya que el número de estados sub-microscópicamente distintos de los dos gases que se consideran estados "mixtos" es muchísimo mayor que el número de los sub-microscópicamente distintos que se consideran "no mixtos".

Supóngase que tenemos n moléculas de un gas en la parte A del recipiente y n' moléculas del otro gas en la parte B del recipiente. Dividamos cada compartimiento en m células pequeñas. Entonces, considerando a cada molécula diferenciable de todas las demás, tenemos $m^n \cdot m^{n'}$ configuraciones distintas. Cuando se quita la barrera hay $2m$ posiciones distintas disponibles para cada molécula, y, por tanto, el número de configuraciones distintas se eleva a $(2m)^n \cdot (2m)^{n'}$, que es mayor mediante un factor de $2^n \cdot 2^{n'}$. No es de extrañar que los estados mixtos de los gases sean incomparablemente más comunes que lo son los estados no mixtos. Sin embargo, esto no significa que *una* configuración submicroscópicamente "mixta" diferenciable cualquiera sea en absoluto más probable que *una* configuración "no mixta" cualquiera.

En la mecánica estadística el estado submicroscópico de un gas depende no solamente de las posiciones de las moléculas sino también de su impulso. (A veces, su impulso angular de rotación y otros parámetros también tendrán que ser discutidos, pero aquí discutiremos, únicamente, su impulso de translación. Considérense n moléculas. Podemos imaginar un espacio de seis dimensiones dando a la molécula k-ava una posición $(x_k, y_k, z_k, p_{xk}, p_{yk}, p_{zk})$, donde (x_k, y_k, z_k) son sus coordenadas espaciales y (p_{xk}, p_{yk}, p_{zk}) los componentes de su impulso. Entonces, en lugar de considerar n moléculas en el espacio de seis dimensiones, podemos considerar una partícula representativa en un espacio de $6n$ dimensiones. Esta útil ficción se denomina espacio de una fase. El principio de conservación de la energía implica que la partícula tratará una trayectoria en una hiper-superficie de $6n-1$ dimensiones en el espacio de la fase. Supóngase que dividimos esta hiper-superficie en un gran número de célu-

las arbitrariamente pequeñas. Entonces, con la excepción de las condiciones iniciales la partícula representativa pasará, finalmente, a través de todas estas células. Este es el "teorema ergódico". Un caso de condiciones iniciales excepcionales en el que este teorema no sería válido es aquél en que todas las moléculas están en un recipiente rectangular desde cuyas paredes se reflejan de modo perfectamente elástico, y todos sus movimientos son paralelos a los lados de la caja, y nunca entran en colisión entre sí. Tal caso no se daría en la naturaleza.

La probabilidad de un estado en que todas las moléculas de un gas tengan las posiciones $q_1, q_2, \ldots, q_n$ y las velocidades $v_1, v_2, \ldots, v_n$, respectivamente, puede demostrarse que es igual a la probabilidad de las posiciones $q_1, q_2, \ldots, q_n$ (igual que antes) junto con las velocidades opuestas $-v_1, -v_2, \ldots, -v_n$. Esto parece presentar la consecuencia paradójica, sobre la cual llamó la atención J. Loschmidt [5], de que los procesos mixtos resultarían tan comunes en un sistema termodinámicamente cerrado y finito como lo son los procesos no mixtos. Una solución parcial a esta paradoja radica en pensar que cualesquiera sistemas pequeños cerrados (o casi cerrados) que observemos no continúan siendo sistemas cerrados (o casi cerrados) durante mucho tiempo, y por tanto los teoremas relativos a lo que ocurre en un intervalo muy largo de tiempo no pueden aplicarse. Con todo, queda el problema, que consideraremos posteriormente, respecto a por qué si consideramos un gran conjunto de sistemas durante un breve tiempo (en lugar de un sistema durante mucho tiempo) no nos encontramos con que los procesos mixtos y no mixtos son igualmente comunes.

Se puede arrojar más luz sobre la paradoja de Loschmidt si adoptamos un determinado presupuesto contrario a los hechos. Supongamos que el universo es un sistema finito cerrado de partículas newtonianas perfectamente elásticas y que ni se está expandiendo ni se está contrayendo. Podemos asegurarnos de su cerradez y finitud presuponiendo que estas partículas existen en un espacio elíptico finito pero no ligado [5]. En tal caso la paradoja de Loschmidt sería realizada, y cualesquiera evoluciones de este universo serían seguidas finalmente de precisamente las vueltas hacia atrás inversas. (O, precisando más, en el sentido de que por pequeñas que fuesen las células en que subdividiésemos el espacio de la fase el "punto representativo" del sistema pasaría finalmente a través de la misma

---

[5] S. Toulmin en su colaboración en *Metaphysical Beliefs* [6], ha hecho una objeción, que parece ser injustificada, al tratamiento del universo como un sistema termodinámico cerrado. Afirma que un sistema termodinámicamente cerrado es aquel que está aislado de las influencias exteriores, pero que no puede haber exterior en el caso del universo. La argumentación de Toulmin parece depender en gran medida de su alegación de que *carece de significado* hablar de los contornos del universo. En mi opinión hablar de los contornos del universo es contradictorio más que carente de significado. Puesto que la negación de una contradicción es una verdad, podemos decir, con verdad, que el universo no tiene contornos. Si definimos "el sistema S está aislado de sus contornos", como "no existen contornos de S que influyan casualmente en S" entonces, de modo trivial, S está aislado cuando no tiene ningún contorno en absoluto.

secuencia de células en el espacio de la fase en orden inverso: "Finalmente" se refiere aquí, por supuesto, a un tiempo inconcebiblemente largo, incluso comparado con el período completo de nuestro sistema galáctico.) Por ejemplo, si el nuestro fuese un universo finito, no expansivo, del tipo aludido en nuestra suposición, personas exactamente iguales a nosotros, en el futuro lejano, estarían realizando exactamente las mismas acciones, y recibiendo exactamente los mismos estímulos sensoriales, excepto hacia atrás en el tiempo [6]. Merece la pena señalar, sin embargo, que tales personas no se sentirían en absoluto diferentes. Puesto que, de acuerdo con nuestra dirección de lo anterior y lo posterior, las huellas de la memoria en los cerebros, en la era del tiempo inverso, se darían antes de los acontecimientos de los que eran huellas, su concepto de lo anterior y lo posterior también sería lo inverso del nuestro, y para ellos el tiempo parecería ir en la dirección inversa de la que parece ir para nosotros. Se daría una completa simetría de imagen de espejo no sólo en el universo en su conjunto, sino, también, en la experiencia subjetiva: a la gente, en una era de tiempo invertido le seguiría pareciendo que la memoria seguía al acontecimiento recordado, y que la herida seguía dándose después del golpe [7].

Con respecto a nuestro universo real, no podemos presuponer que sea finito (aunque pudiera serlo) ni que *no* se esté expandiendo, puesto que los desplazamientos hacia el rojo observadas en galaxias lejanas sugieren con bastante base que el universo se *está* expandiendo. La termodinámica cosmológica es, por tanto, una materia difícil y cuestionable. Consideremos la cuestión de la aparente asimetría temporal de esa región espacio-temporal del universo en la que nos encontramos y que podemos observar con facilidad. ¿Cómo es que, dentro de esta región al menos, parece haber una asimetría con relación a las huellas, pudiendo haber, por ejemplo, fotografías o cintas magnetofónicas de acontecimientos anteriores pero no de los posteriores? Consideremos las características termodinámicas de tales cosas.

Supóngase que contamos con un dado inestable, siendo las probabilidades para sus caras 1, 2, 3, 4, 5, 6, p_1, p_2, p_3, p_4, p_5, p_6, respectivamente. Evidentemente $p_1+p_2+p_3+p_4+p_5+p_6=1$. Supongamos que conocemos estas probabilidades y que se nos dice que el dado será tirado en una determinada ocasión. ¿Cuál es nuestra falta de seguridad con relación a los resultados? Supóngase que $p_6=0{,}99$ y que $p_1=p_2=p_3=p_4=p_5=0{,}002$. En este caso podemos estar casi seguros que saldrá un 6. Es poca la inseguridad con relación a los resultados que queda eliminada mediante la tirada real, puesto que en la mayoría de los casos el resultado de la tirada confirmará simplemente nuestros cálculos. Desde un punto de vista intuitivo

---

[6] De hecho un número infinito de veces, tanto en la misma dirección temporal como en la dirección inversa.

[7] F. H. Bradley consideraba que en un universo en el que la dirección del tiempo estuviese invertida, "La muerte llegaría antes del nacimiento, el golpe seguiría a la herida, y todo debería parecer irracional" [7].

parece plausible afirmar que la máxima inseguridad que viene a ser disipada por una tirada real se da cuando $p_1=p_2=p_3=p_4=p_5=p_6=\frac{1}{6}$. Considérese, pues, esta definición de la "información" (debida a C. E. Shannon [8]) que ha de ser proporcionada por una tirada real

$$H=-\sum_{1}^{6} p_i \log p_i$$

En el caso en el que todas las p_i sean iguales a $\frac{1}{6}$ esto se convierte en $-\frac{1}{6}\log\frac{1}{6}$. (Obsérvese que $\log\frac{1}{6}$ es negativo y por tanto $-\log\frac{1}{6}$ es positivo.) Cuando no hay inseguridad respecto al resultado $H=0$. Cuando los p_i son muy desiguales H se acerca a 0. Por ejemplo, conforme p_6 se aproxima a 1, la cantidad H se aproxima a 0, y también podemos decir que $H=0$ en el caso en que $p_6=1$ y $p_1=p_2=p_3=p_4=p_5=0$, aun cuando $\sum_{1}^{6} p_i \log p_i$ esté indeterminado en este caso. Podemos sospechar ya que H proporciona una buena medida de algo semejante a la información, ya que en el caso en que $p_6=1$, no hay ninguna inseguridad que haya de ser disipada, y, por tanto, no se obtiene ninguna información mediante una tirada del dado. En general si las probabilidades de n acontecimientos alternativos son $p_1, p_2, \ldots, p_n$, entonces la información que se consigue observando cuál de ellos ocurre realmente viene dada por

$$H=-\sum_{i=1}^{n} p_i \log p_i$$

Recuérdese ahora la noción de espacio de la fase en la mecánica estadística y hágase que p_i sea la probabilidad de que la partícula representativa está en la célula i-ava del espacio de la fase. Puede verse que H, tal como se definió anteriormente, se convierte en máxima cuando las p_i son todas iguales. Es este también el caso en que la entropía es máxima. De hecho H, tal como se definió anteriormente, es exactamente menos una vez la cantidad H de Boltzmann, y es equivalente a la entropía tal como es definida por Boltzmann. El teorema H de Boltzmann afirma que en un sistema cerrado su H disminuye, lo cual viene a ser casi lo mismo que decir que la entropía aumenta. Desde luego, a veces la entropía es definida simplemente como $\sum_{i=1}^{n} p_i \log p_i$ (o, mejor, haciendo las células del espacio de la fase infinitamente pequeñas, como $\int F \log F d\omega$, siendo F una función que determina la probabilidad de encontrar a una molécula muy cerca de un punto determinado del espacio de la fase).

Esta conexión entre la información, tal como se define en la teoría de la información, y la entropía, tal como se define en la mecánica estadís-

tica, ha dado lugar a una de las atrayentes ideas de la ciencia reciente (aunque ha de ser tratada con cautela), y se hace inmediatamente más plausible que el contenido informativo de cosas tales como fotografías o cintas magnetofónicas tenga algo que ver con su escasa entropía como sistemas físicos.

Puede ser que esto sea así, aun cuando se dé el caso de que el término "información" en la teoría de la información sea un término técnico, que se utiliza en un sentido bastante distinto del ordinario. Examinemos brevemente algunas de las diferencias.

Si llevamos los logaritmos de la fórmula $H = -\sum_{i=1}^{n} p_i \log p_i$ a la base dos, nuestra unidad de información (denominada "dib", abreviatura de "dígito binario") *, corresponde al caso en que $H = -(\frac{1}{2} \log \frac{1}{2} + \frac{1}{2} \log \frac{1}{2})$, ya que esto es igual a uno cuando nuestros logaritmos son llevados a la base dos.

Supóngase ahora que contamos (como es posible, por supuesto) con un alfabeto que consta únicamente de ceros y de unos. Hagamos que sea un hecho empírico el que la frecuencia de ceros en un trozo largo de discurso sea $\frac{1}{2}$, y de modo semejante para los unos. Hagamos también que se dé el caso, no probable en ningún alfabeto real, de que la probabilidad de que se dé 0 ó 1 sea completamente independiente de los símbolos que los hayan precedido. (Un caso extremo, que muestra que este tipo de presupuesto no es posible en inglés, es el de que la probabilidad de "*u*" siguiendo a "*q*" se acerque mucho a 1, mientras que la frecuencia de las *u*s en cualquier trozo largo de discurso es bastante pequeña, y nada cercana a 1.) La teoría de la información puede, desde luego, tratar casos en que la probabilidad de que aparezca un símbolo depende de las probabilidades de aquéllos que le han precedido (procesos Markoff), pero aquí, para mayor simplicidad matemática, estamos considerando una secuencia de ceros y de unos, que son tales que la probabilidad de que aparezca un 0 o un 1 es siempre $\frac{1}{2}$, incluso con relación a cualesquiera símbolos hayan aparecido ya. Debo insistir en que esto se utiliza aquí únicamente como un mecanismo expositivo simplificador, ya que en cualquier lenguaje real, las reglas de ortografía y sintaxis, así como también, con toda probabilidad, las semánticas también, impedirían la probabilidad de que el que apareciese un símbolo determinado fuese independiente de su lugar en la secuencia de símbolos. (Cualquier lenguaje real debe contener una buena proporción de "redundancias", en el sentido dado a esta palabra en la teoría de la información.)

Con el presupuesto mencionado, pues, considérese una secuencia de n ceros y unos que sea transmitida por algún sistema de señales. Este proporciona n "dibs" de información, en el sentido de que n "dibs" de inseguridad respecto a lo que será enviado son resueltas cuando se recibe el

* En el original inglés "bit", como abreviatura de "binary digit". *(Nota de la Traductora.)*

mensaje. Supongamos ahora que el emisor y el receptor utilizan el código binario mencionado anteriormente para enviar información en el sentido *ordinario* de la palabra, es decir, lo que denominaremos "información semántica", del modo siguiente: Supóngase que hay 100 partidos de "squash" [8] que van a ser jugados, siempre entre equipos bien constituidos, de forma que en todos los casos el receptor del mensaje esté del todo inseguro con respecto al resultado, en el sentido de que se sienta inclinado a conceder una probabilidad de $\frac{1}{2}$ a que gane cada uno de los equipos. El emisor y el receptor del mensaje tienen ambos una lista numerada de las competiciones, y lo convenido es que si el dígito n-avo transmitido es un "0" esto significa que el equipo de casa ha ganado, y si ese un "1" significa que el equipo forastero ha ganado. En este caso podemos igualar la medida de la "información semántica" con la información de la "teoría de la información": la inseguridad con relación a qué secuencia de ceros y de unos será enviada es n "dibs", y lo mismo lo es la inseguridad respecto a los resultados de las n competiciones. En general, sin embargo, no podemos igualar la información de la teoría semántica con la información de la teoría de la información. Esto se debe, en parte, a que, en general, no hay una única correlación entre la probabilidad, pongamos por caso, de que sea enviada la palabra "azul" y la probabilidad de que sea azul la cosa del mundo acerca de la cual se predica tal cosa. Cuestiones de redundancia también complican este punto. Pero los motivos centrales por los que no se da, generalmente, una simple relación entre la información de la "teoría de la información" y la información "semántica" son los dos siguientes: 1) No es fácil dar una definición de la información semántica, excepto desde determinados supuestos artificiales, que presuponen un lenguaje cuyas proposiciones son o bien atómicas o de lo contrario sus funciones de verdad [9]. 2) Supóngase que en un caso tratable particular, tal como el mencionado anteriormente de los partidos de "squash", convenimos que el dígito $(m+1)$-avo proporciona el resultado del partido m-avo, y que el primer dígito se reserva para informarnos de si han ganado *todos* los equipos de casa o no. Si todos los equipos de casa efectivamente ganasen, entonces el primer dígito nos daría n "dibs" de información semántica, y sólo un "dib" de información de la "teoría de la información". Porque la información de la teoría de la información fue calculada de acuerdo con el presupuesto de que la probabilidad de cero o de uno era siempre de $\frac{1}{2}$, aun cuando, una vez que conocemos la regla semántica, podemos ver que esto no es cierto con relación al primer dígito.

A pesar de lo anterior, sigue siendo, con toda seguridad, verdadero que en cualquier sistema de señales, si no se envía información de la teoría de la información —si el canal emite únicamente "ruidos"— entonces no se envía información semántica alguna por el canal. (Por supuesto

[8] He elegido el "squash" en lugar de cualquier otro juego por la simplicidad para evitar la posibilidad de ligas.

[9] Una definición de la información semántica relativa a un lenguaje bastante simple ha sido proporcionada por Y. Bar-Hillel y R. Carnap [9].

que esto puede proporcionarnos *indirectamente* información con relación al canal, e incluso con relación al emisor. Si observamos la caída de bombas de artillería cerca de una estación de radio que esté a punto de empezar a operar, y a continuación no oímos en nuestros audífonos más que diversos crujidos, sospechamos que la estación de radio ha sido derribada.)

En la teoría de la información nos ocupamos, generalmente, de lo que Reichenbach ha denominado "macroentropía" [10]. En la mecánica estadística, que podemos denominar "microestadística", nos ocupamos de la distribución estadística de las posiciones e impulsos de las moléculas, pero en la macroestadística nos ocupamos de la distribución estadística de las propiedades de objetos macroscópicos. Por ejemplo, los objetos de nuestra macroestadística pueden ser las arenas de una playa. Por lo que sabemos de probabilidades, es mucho más probable un trozo de playa bastante liso que un trozo de arena que contenga huellas de pisadas. Es decir, salvo en el supuesto de la interacción de un hombre que caminase por la playa con la arena de la playa, un surco en la playa con la figura de una huella de una pisada es muy improbable. El viento y el tiempo atmosférico podrían, por casualidad, haber cavado tal figura, pero tal cosa no es en absoluto previsible. Por esta razón, la presencia de una huella de pisadas es un indicador seguro de la interacción de la playa con un sistema exterior. Podemos decir que la macroentropía de la playa ha sido disminuida. Es la disminución de la macroentropía lo que hace de la huella de pisada un indicador seguro, y no la disminución de la microentropía, aun cuando en este caso sea plausible suponer que la microentropía ha sido disminuida, desde luego. (La microentropía del sistema total, playa más caminante, habrá aumentado, por supuesto, y cualquier disminución de la microentropía de la playa será a costa de un aumento más que compensatorio de la entropía del cuerpo del caminante. Hasta que pueda adquirir nuevas energías en su próxima comida, cada nuevo paso reduce la energía que tiene acumulada y da como resultado un aumento de la entropía de su cuerpo.)

Para ver claramente qué es la disminución de la macroentropía, no necesariamente de la microentropía, lo que convierte a un sistema en un registro, considérese el caso de la reproducción fotográfica de una página impresa. Cuando un fotón da contra una placa fotográfica la ennegrece ligeramente. Esto es debido a que la energía de la luz desprende la unión química en una molécula de bromido de plata en la emulsión, llevando a un aumento de la entropía. No hace falta de que nos ocupemos, sin embargo, de todos los detalles químicos (por ejemplo, de si se da o no se da un aumento de entropía en las regiones ennegrecidas de la placa, después del proceso químico posterior del fijado) porque si (supongamos) el ennegrecimiento corresponde a un cambio de entropía en una dirección, entonces el enblanquecimiento corresponde a un cambio de entropía en la otra dirección. Ahora bien, puesto que una página de letras negras sobre blanco proporciona la misma información que blanco sobre negro (tiene

la misma macroentropía) es del todo evidente que una disminución de la macroentropía podría corresponder tanto a un aumento de la microentropía como a una disminución de la misma.

El punto anterior puede aclararse mediante la figura 14. Aquí tenemos una secuencia de áreas de elevada microentropía (regiones desordenadas) intercaladas con áreas ocasionales de escasa microentropía (regiones muy ordenadas), que funcionan, quizá, como los puntos y rayas de un código de Morse. Esto no puede transmitir una señal ni mejor ni peor que el caso de la figura 15, que es igual al de la figura 14, excepto en que cuando hay regiones de entropía elevada en una figura, entonces hay regiones de escasa entropía en la otra, y viceversa. Las regiones ordenadas y desordenadas pueden ser las regiones magnéticas y no magnéticas de una cinta, por ejemplo.

FIGURA 14.

FIGURA 15.

Con todo, probablemente sigue siendo cierto que en casi todos los casos de registros *naturales* (por ejemplo, pisadas, fósiles) los sistemas de escasa macroentropía, de hecho, también son sistemas de escasa microentropía. Ya que la norma es una tendencia hacia la microentropía elevada, y es normalmente la microentropía lo que hace que un sistema se salga de sus contornos como prueba de una interacción entre el sistema que la contiene y algún sistema exterior (por ejemplo, entre la playa y un caminante).

Una pisada en la arena tenderá a aumentar su entropía hasta que su entropía sea una vez más, una característica del sistema más amplio del que es parte. En algunos casos, sin embargo, un subsistema de escasa entropía permanecerá en equilibrio hasta el momento en que pierda su individualidad como sistema. Considérese una cinta magnética que no pierde prácticamente nada de su magnetismo hasta un momento tal en que la cinta es destruida. Pero, ya bien posea una entropía incrementada, o ya bien permanezca en equilibrio termodinámico, el subsistema en cuestión permanece en un estado de escasa entropía si lo comparamos con lo que habría de esperarse de compararlo con sus contornos. (Es necesario expresar la cuestión de este modo un tanto molesto: no sirve decir "su entropía es inferior a la de sus contornos", puesto que la entropía depende de la masa, y un subsistema pequeño, por tanto, tiene, generalmente, menos entropía que el resto del sistema mayor del cual forma parte.)

Un sistema que sea temporalmente (o, como Grünbaum ha señalado [11], *indefinidamente,* si está en equilibrio) de entropía más escasa de lo que debería esperarse cuando lo comparamos con sus contornos es el tipo de sistema que Reichenbach [12] denominó sistema *derivado* (*branch* system). Riechenbach consideraba a tales sistemas derivados de la curva de la entropía del universo como un todo. El mantuvo la hipótesis de que la dirección de entropías más escasas a más elevadas, de los subsistemas, era en casi todos los casos la misma que la dirección de la entropía más escasa a más elevada del universo completo. La formulación que daba fallaba al no dar cuenta de la cuestión discutida en la sentencia entre paréntesis al final del parágrafo precedente. A pesar de esto, se establece fácilmente del modo antes mencionado. Pero sus observaciones presentan otro defecto más. ¿Por qué incluir en ella la entropía del universo completo? Sabemos muy poco con relación a la entropía del universo completo y, además, no es evidente que la entropía pueda ser definida de un modo significativo para el universo como un todo, si el universo es un sistema infinito, como se supone en algunas hipótesis cosmológicas. Esta referencia a la curva de la entropía del universo no es esencial, sin embargo, y Grünbaum ha sustituido la hipótesis de Reichenbach por la hipótesis, que constituye una generalización de lo que observamos que se da *de facto* a nuestro alrededor, de que casi todos los sistemas tales pasan de la entropía más escasa a la más elevada (a menos que estén en equilibrio) en la misma dirección unos que otros. Esto evita la cuestionable referencia al universo como un todo. Podemos, pues, definir la dirección de lo anterior a lo posterior como la dirección de la entropía más escasa a la más elevada de la mayoría de los sistemas derivados.

Si definimos la dirección de lo anterior a lo posterior, la proposición de que si todos los sistemas derivados aumentan su entropía se hace verdadera por definición, pero, con todo, no ocurre esto de un modo trivial. Depende del hecho de que casi todos los sistemas derivados van en la misma dirección. Más aún, deja lugar para que la afirmación de que un sistema aumenta su entropía sea empírica.

Desde luego, esta definición de la dirección de lo anterior a lo posterior no es un pre-requisito para aprender nuestro uso ordinario de palabras tales como "anterior" y "posterior". Es evidente que los niños aprenden palabras tales como "anterior" y "posterior" mucho antes de que conozcan el significado de palabras tales como "entropía". Esto se debe a que probablemente aprenden estas palabras mediante algo semejante a una definición ostensiva. En una situación ideal podemos imaginar que cogemos a un niño y le enseñamos una naranja durante breves instantes, luego una manzana, y le decimos cosas tales como "naranja anterior, manzana posterior". Sin embargo, si nuestra hipótesis es correcta, respecto a que las huellas de la memoria en el cerebro son subsistemas de entropía temporalmente más escasa de lo que es característico en sus contornos, entonces la posibilidad física de tal aprendizaje ostensivo de las palabras "anterior" y "posterior" implica que sea correcta la presuposición de nuestra defi-

nición física de la dirección del tiempo, a saber, la presuposición de que casi todos los sistemas derivados van en la misma dirección. Se sigue de lo anterior que los intentos de tratar la relación "anterior-posterior" como indefinible (con la base, desde luego cierta, de que los niños no aprenden a utilizar las palabras "anterior" y "posterior" mediante definiciones verbales) ocultan la base física del sentido de la dirección de esta relación [10].

Es cierto, desde luego, que podríamos definir una relación *ordenadora* de "anterior" y "posterior" incluso en un universo temporalmente simétrico. Podríamos fijarnos en dos acontecimientos arbitrarios E_1 y E_2 y hacer las siguientes estipulaciones: Si un acontecimiento A está entre E_1 y E_2 entonces es posterior a E_1 y anterior a E_2. Si no está entre E_1 y E_2, sino que E_2 está entre E_1 y A, entonces A es posterior tanto a E_1 como a E_2. Si A no está entre E_1 y E_2, sino que E_1 está entre A y E_2, entonces A es anterior tanto a E_1 como a E_2. (Aquí hemos adoptado el presupuesto de que los sucesos pueden ser ordenados [11] en un tiempo topológicamente abierto: al igual que una línea que sea infinita en ambas direcciones. Para un tiempo cerrado, tal como es concebible que pudiera ser postulado por algunos cosmólogos, se necesitaría una definición bastante más complicada.) La cuestión es, sin embargo, que tal definición de "anterior" y "posterior" para un universo temporalmente simétrico sería arbitraria: no importaría que eligiésemos nuestros acontecimientos E_1 y E_2 de modo que la dirección del tiempo positivo fuese en el sentido AB (siendo A y B dos acontecimientos definidos) o de modo tal que fuese en el sentido BA. Estaríamos dando una dirección arbitraria al tiempo, al igual que pudiéramos trazar arbitrariamente un eje espacial ya bien en la dirección del sur al norte o en la dirección del norte al sur.

La definición de la dirección de lo anterior a lo posterior en términos de la dirección del aumento de la entropía de la mayoría de los sistemas derivados no es arbitraria, como sería la hecha en términos de la dirección de uno a otro de dos acontecimientos elegidos arbitrariamente. La dirección se corresponde efectivamente con algo en la constitución del universo, o al menos con aquella parte de él con la que estamos familiarizados.

Podríamos preguntarnos qué tipo de hecho es el de que la gran mayoría de los sistemas derivados de nuestra experiencia vayan de la entropía escasa a la entropía elevada (a menos que estén en equilibrio) en la misma dirección temporal. Grünbaum lo considera como nomológicamente contingente, o, en otras palabras, como algo que no se sigue de las leyes de la naturaleza únicamente. Es exactamente un hecho [15] que parece ser sugerido por los ejemplos de los sistemas derivados que observa-

---

[10] Por consiguiente, me parece a mí que debe de haber algo que podría ser, por lo menos, desorientador en la afirmación de Max Black en su ensayo "The Direction of Time" [13] de que el que un acontecimiento sea anterior a otro depende de estos acontecimientos únicamente. Sin embargo, no estoy seguro de estar en desacuerdo con Black realmente, especialmente en vista de su cauto parágrafo final.

[11] Grünbaum denomina a la relación ordenadora de entre dos (betweenness) para un tiempo topológicamente abierto, "o-betweenness" [14].

mos en el universo que nos rodea. Este hecho nomológicamente contingente puede ser expresado como un postulado [16] relativo a la azarosidad de la distribución de los micro-estados correspondientes a un determinado macro-estado. Supóngase que el macro-estado que nos ocupa consiste en leche por encima del té en una taza. Esto es ejemplificado cada vez que alguien echa leche al té. Los diversos micro-estados correspondientes a los estados macroscópicamente indiferenciables de "leche encima del té" se distribuyen al azar, de modo que prácticamente ninguno de ellos es uno de esos raros estados que no llevarían rápidamente a una mezcla de la leche con el té.

Supóngase, sin embargo, que existen dos porciones distintas del universo para cada una de las cuales es válido el postulado de la azarosidad. Supongamos que "la dirección del tiempo", es decir, la dirección en la cual la mayoría de los sistemas derivados incrementan su entropía, es opuesta en los dos sistemas. Si las dos partes del universo tuviesen interacción sería difícil que imaginásemos la situación, ya que esto dificultaría nuestro presupuesto de la azarosidad. Consideremos, sin embargo, que estas dos partes del universo apenas tuviesen interacción alguna, sino que se diese una interacción casual esporádica. Aquí tendríamos un caso semejante al imaginado por Norbert Wiener [17]. Supóngase que se arroja una piedra por un *A* de la primera parte a la otra parte, donde, según *A*, *B* la recoge. Pero, puesto que *B* considera la dirección del tiempo en el sentido opuesto al de *A* se considerará a sí mismo tirando la piedra y dirá que es él quien la tira y *A* el que la recoge.

Supóngase que encontrásemos un ejemplar de la Biblia en el césped exterior a nuestra casa. Conforme lo miramos vemos que de repente sube por los aires y se pierde en las profundidades del espacio. Tal acontecimiento es físicamente [12] posible: todas las moléculas de la Tierra podrían coincidir en ascender en el mismo momento, impartiendo así su energía cinética a la Biblia y lanzándola al espacio. Sin embargo, tal suceso sería tan improbable que sería preferible la hipótesis de que la Biblia había sido arrojada a nuestro mundo por algún misionero de un mundo con un tiempo invertido.

La mencionada historia fantástica sugiere, por lo menos, que la hipótesis de regiones del universo temporalmente invertidas pero simultáneas no tiene por qué carecer de posibilidades de una comprobación empírica. Con todo, la posibilidad de tal caso parece ser casi infinitamente improbable. Exigiría la división del espacio de la fase para todo el universo completo en dos partes distintas, lo cual significaría que la entropía del universo combinado sería mucho más elevada de lo que pudiera parecer probable en absoluto. (En contra del caso de los aumentos de entropía, a

[12] En realidad esto no es correcto, a menos que el suceso tenga lugar en un planeta de masa muy baja y sin atmósfera. De otro modo no habría energía suficiente en la parte del césped bajo la Biblia. Pero tales consideraciones no importan aquí, cuando nuestro ejemplo es evidentemente imaginario, y si lo deseamos podemos adoptar presupuestos igualmente imaginarios respecto a la masa de nuestro planeta.

través del tiempo infinito, de un universo finito pero no en expansión y no ligado, donde, como hemos visto, los períodos de aumento y disminución de la entropía, en relación con una dirección del tiempo positivo elegida arbitrariamente, serían igualmente comunes.)

Hemos estado intentando relacionar la asimetría temporal de al menos una amplia región cósmica del universo que nos rodea, con consideraciones relativas a la entropía y a la mecánica estadística. Veamos cómo estas consideraciones estadísticas enlazan, si es que lo hacen, con otras dos características que algunos escritores han considerado que eran la base de la asimetría temporal del mundo. Son: 1) el principio de las ondas retardadas, y 2) el hecho de la expansión del universo.

El principio de las ondas retardadas

Observamos anteriormente que las ecuaciones del electromagnetismo son temporalmente simétricas, en el sentido de que si $f(t)$ es una solución a ellas, también lo es $f(-t)$. Por ejemplo, una onda de luz esférica que surgiese de un punto de origen no es ni más ni menos consistente con las ecuaciones de Maxwell que una onda de luz esférica que proceda del infinito y se contraiga hacia un punto de extinción. El primer caso es el de una onda retardada, y el segundo el de una onda avanzada. Cuando calculamos el potencial eléctrico ϕ en un punto P y en un momento de tiempo t debido a una carga variante ρ a una distancia r de P, lo hacemos basándonos en el valor de ρ en el momento de tiempo $t-\frac{r}{c}$ (es decir, en el momento en que una señal luminosa que parta de la carga llegue a P en el momento de tiempo t) y no en el momento de tiempo $t+\frac{r}{c}$, que proporcionaría igualmente una solución a las ecuaciones de Maxwell [18]. Esto se dice que es tomar el potencial *retardado*. Si calculásemos el potencial sobre la base de ρ en $t+\frac{r}{c}$ esto significaría tomar el potencial avanzado. La solución "retardada" nos proporciona el caso de una onda emitida desde la carga variante, mientras que la solución "avanzada" nos proporciona una onda que viene a la carga ρ y le hace variar. El principio de las ondas retardadas es exactamente éste: nos encontramos con montones de casos en los que evidentemente las ondas son emitidas desde puntos de origen y marchan al infinito, pero ninguno en que las ondas procedan del infinito y vayan a un punto de extinción. Hay, por supuesto, ondas que van a puntos de extinción y que pueden ser consideradas, cuando discutimos antenas receptoras de radio, por ejemplo, como ondas "avanzadas", pero todas éstas son en sí mismas también ondas "retardadas" procedentes de algún punto de origen, por ejemplo alguna antena transmisora.

¿Es este principio de las ondas retardadas independiente del principio

de aumento de la entropía? Si se diese el caso de que una onda contraída esféricamente estuviese disminuyendo su entropía, mientras que una onda expandida esféricamente estuviese aumentando su entropía, entonces el principio de las ondas retardadas sería simplemente un caso especial del criterio de entropía de la dirección del tiempo. Sin embargo, la definición habitual de entropía en la mecánica estadística depende de una finitud de espacio de la fase, y no puede ser aplicada al caso de un sistema espacialmente infinito. La mecánica estadística hace uso de la hipótesis cuasi-ergódica de que la partícula representativa en el espacio de la fase pasa aproximándose arbitrariamente a todos los puntos del espacio de la fase, y es difícil ver cómo esto pudiera parecer plausible en el caso de un sistema infinito. Queda por resolver si es posible una sustitución adecuada de la hipótesis cuasi-ergódica y tal vez una redefinición plausible de la entropía, que hiciese que la mecánica estadística pudiese abarcar los casos de procesos en el espacio infinito, y que mostrase a las ondas retardadas como casos de entropía en aumento, y a las ondas adelantadas como casos de entropía en aumento, y a las ondas adelantadas como casos de entropía en disminución.

O. Costa de Beauregard [19] ha mantenido la equivalencia de los dos principios (el del aumento de la entropía y el de las ondas retardadas) valiéndose de la consideración de los fenómenos de dispersión. Estableciendo a mi manera lo que yo considero que es su argumentación, viene a ser lo siguiente: Supóngase que una onda óptica coherente es dispersada mediante reflexión en una retícula del microscopio o en un espejo. La onda reflejada es incoherente. Si hubiésemos de invertir el proceso (tomar las soluciones avanzadas en lugar de las retardadas) tendríamos el caso muy improbable de que un número de ondas incoherentes produjesen, al reflejarse, ondas coherentes. Por tanto, la cuestión de tomar la solución retardada equivale a no esperar que ocurra lo improbable, es decir, es equivalente al principio de aumento de la entropía. De Beauregard señala que incluso cuando en la mecánica estadística consideramos un número de partículas, éstas deben ser estimadas en términos de la mecánica ondulatoria como fenómenos ondulatorios, y el principio de entropía en aumento resulta ser un fenómeno de dispersión, exactamente igual que el caso óptico, y el principio de aumento de la entropía resulta ser equivalente, una vez más, al principio de las ondas retardadas. Si la posición de Beauregard es correcta, entonces Popper, que basaba la asimetría temporal en la prevalencia en nuestra experiencia de ondas expansivas, en lugar de ondas contraídas, no estaba proporcionando una alternativa distinta del método estadístico respecto al problema de la asimetría temporal [20].

La asimetría temporal y la expansión del universo

Por término medio, las galaxias distantes parecen estar alejándose de nosotros. Esto se deduce del desplazamiento hacia el rojo de ciertas

líneas reconocibles de sus espectros, que es más plausible interpretar como un cambio Doppler debido a la recesión en la línea de la vista. Aproximadamente, la velocidad de recesión es proporcional a la distancia a partir de nosotros. (Esta relación lineal quizá no sea del todo correcta, sin embargo, puesto que observaciones recientes han sugerido una cierta separación de la ley lineal en el caso de galaxias muy distantes.) No debe considerarse, sin embargo, que esta recesión de las galaxias es una recesión a partir de la nuestra en particular. Si existe una expansión general del universo, cada galaxia será el centro de una recesión a partir de ella, siendo la recesión aproximadamente proporcional a la distancia. Un ejemplo, dado una vez por Eddington [21], servirá aquí. El estaba pensando en términos de un espacio esférico finito pero no ligado, análogo tridimensional a la superficie de la esfera. Considérese entonces un globo de goma en que hay anotados un cierto número de puntos, estando repartidos por igual por todo el globo. Conforme se infla el globo, dos puntos cualesquiera se separarán, siendo la velocidad de la recesión proporcional a su distancia. Es evidente que se observaría el mismo fenómeno si se diese una expansión general de un espacio infinito. De momento no nos preocupa, por consiguiente, si nuestra teoría cosmológica postula un universo finito pero no ligado (como el universo de Einstein) o uno infinito (como en las teorías del estado estable).

Si interpretamos el desplazamiento hacia el rojo, pues, como prueba de que el universo se está expandiendo del modo anteriormente mencionado, contamos con un criterio respecto a la dirección de nuestro eje del tiempo. Es decir, la dirección del tiempo positivo es aquella de acuerdo con la cual decimos que el universo se está expandiendo. Si tomásemos la dirección del tiempo en sentido contrario (es decir, en términos vulgares, del futuro al pasado, en lugar del pasado al futuro), diríamos que el universo se está contrayendo. En otras palabras, la dirección del tiempo positivo es aquella de acuerdo con la cual debemos decir que el universo se está expandiendo.

Si tomásemos la dirección del tiempo positivo en sentido opuesto al que la tomamos actualmente, esto daría lugar a una descripción diferente de un modo meramente convencional, de los hechos en cuestión. (Hay otro sentido más importante en el que podríamos hablar de un universo contraído, a saber, aquél en el que tuviéramos que dar cuenta de la observación no de un desplazamiento hacia el rojo sino de un desplazamiento hacia el azul.) Después del cambio meramente convencional de la dirección del tiempo siguen existiendo los desplazamientos hacia el rojo, pero de ellos se habla de la siguiente manera: Cuando invertimos el eje del tiempo, debemos hablar no de un fotón abandonando la galaxia distante y tocando la placa fotográfica de nuestro telescopio, sino del fotón abandonando la placa fotográfica, pasando a través del telescopio y lanzándose al espacio, siendo absorbido finalmente por un átomo de la galaxia distante. Debido al hecho de que (como ya debemos saber ahora) nosotros y la galaxia nos acercamos mutuamente, el fotón aumenta en energía y es

capaz de meterse dentro de un átomo con una absorción de energía que es mayor que la de la emisión del fotón por el átomo de la placa fotográfica. Compárese esto con la idea de una *verdadera* contracción del universo, es decir, una contracción en el sentido de que las galaxias no mostrasen desplazamientos hacia el rojo sino desplazamientos hacia el azul. (Veremos que los desplazamientos hacia el azul presentan problemas físicos: ¡cuando menos, un universo que mostrase desplazamientos hacia el azul probablemente resultaría desagradablemente caluroso!)

Algunos escritores, incluyendo a T. Gold, han tratado de demostrar que la expansión del universo es la asimetría fundamental y que la asimetría termodinámica discutida anteriormente en este capítulo es un mero corolario de la asimetría cosmológica de la expansión. Gold ha proporcionado un llamativo ejemplo con el que apoyar su punto de vista [22]. Este ejemplo debe ser entendido a la luz de una famosa consideración de la cosmología, la paradoja de Olbers.

A principios del siglo diecinueve, el astrónomo Olbers consideró la posibilidad de que las estrellas estuviesen distribuidas por término medio de un modo uniforme por todo el espacio, que, naturalmente, por aquel entonces él presuponía euclidiano. (Hoy en día diríamos "galaxias" o "conglomerados de galaxias" en lugar de "estrellas", pero la argumentación no se vería afectada por el cambio.) Considérese una concha esférica delgada cuyo centro es la Tierra y cuyo radio es r. El número de estrellas de la concha será proporcional a r^2. La intensidad de la luz que nos llega de cada una de las estrellas de la concha será proporcional a $\frac{1}{r^2}$. Por consiguiente la cantidad de luz que nos llega desde cualquier concha distante de un determinado grosor es proporcional a $r^2 \cdot \frac{1}{r^2}$, es decir, a una constante, y por tanto es independiente del radio de la concha. Desde luego, el supuesto de que el espacio es euclidiano no es esencial para la anterior argumentación, ya que si rechazamos este presupuesto, la argumentación implicaría simplemente la sustitución de r^2 y $\frac{1}{r^2}$ por $f(r)$ y $\frac{1}{f(r)}$ en el argumento mencionado [23]. Parecería seguirse, pues, integrando a r del cero al infinito, que la suma de luz que nos llega será infinita en su cantidad[13]. Incluso de noche, el firmamento sería infinitamente luminoso. Esta conclusión, desde luego, sí que necesita ser modificada a la luz de la consideración de que las estrellas muy distantes estarían oscurecidas por las estrellas más próximas, al igual que los árboles muy distantes de un bosque no pueden verse debido a que están ocultos por los árboles

[13] Esto es lo más fácil de imaginar de acuerdo con el presupuesto de un espacio abierto topológicamente, euclidiano o hiperbólico. Con todo, la integración de cero al infinito podría hacerse todavía en el supuesto de que el espacio sea esférico, puesto que la luz podría haber viajado una y otra vez alrededor del universo cualquier número de veces, presuponiendo una teoría del estado-estable. Sin embargo, el presupuesto de un estado-estable no es, probablemente, compatible con el de un universo esférico.

más cercanos. Sin embargo, la luz de las estrellas muy lejanas calentaría a las estrellas más cercanas y de este modo aumentaría la radiación que ellas producirían por su propia cuenta. Parece evidente que cuando se toman en cuenta estas consideraciones resulta que el firmamento nocturno, aunque no infinitamente luminoso, debiera ser, desde luego, muy luminoso. Ni qué decir tiene que presuponiendo un equilibrio termodinámico a gran escala, es evidente que no habría diferencia entre la luminosidad manifiesta de las estrellas más cercanas y los espacios de firmamento intermedios.

La conclusión del argumento de Olbers es, desde luego, paradójica. Lejos de sentir el intenso calor del firmamento nos sentimos (afortunadamente) bastante frescos, y cuando contemplamos el firmamento de noche vemos no una llamarada de luz sino una suave oscuridad. El conocido hecho de que el firmamento esté oscuro por la noche muestra que debe haber algo equivocado en uno de los presupuestos de Olbers. Uno de los presupuestos ocultos que Olbers tomó en consideración es el de que si las estrellas son en todas partes de la misma luminosidad intrínseca entonces la luz que nos llegue desde una estrella determinada será normalmente proporcional al cuadrado inverso de la distancia que guarde con nosotros. (No es necesario que cuestionemos aquí este presupuesto del espacio euclidiano.) Sin embargo, debido al hecho del cambio hacia el rojo, que implica una pérdida de energía en la cantidad de radiación que nos llega, esta simple aplicación de la ley del cuadrado inverso no es válida. A causa del desplazamiento hacia el rojo, las conchas distantes de un determinado grosor contribuirán con menos radiación a la totalidad de la que nos llegue a nosotros que conchas más cercanas. Desde luego, las galaxias muy distantes tendrían una velocidad de recesión mayor que la velocidad de la luz [14], y no podría llegarnos ninguna radiación de ellas.

La expansión del universo, por consiguiente, explica por qué el firmamento está oscuro por la noche y por qué el universo en su conjunto constituye un "punto de extinción no reflexivo para la radiación". Gold sugiere que éste es el hecho último detrás de la "flecha del tiempo". En la primera parte de este capítulo considerábamos la cuestión desde el punto de vista de la termodinámica, y de las huellas y registros como sistemas derivados cuya entropía es temporalmente inferior a la normal. (El que la microentropía o la macroentropía sean cuestionables aquí es una nueva complicación, pero en la mayoría de los sistemas naturales, en oposición a los mecanismos de señales artificiales, podemos presuponer plausiblemente que la macroentropía escasa implica una escasa microentropía.) Gold sugiere que observemos más profundamente y, en efecto, su posición, es la de que es la expansión del universo la que explica la formación de sistemas derivados. Completamente aparte de cualquier interés

[14] No se está en conflicto aquí con la teoría especial de la relatividad, puesto que estamos en el ámbito de la teoría general y la cosmología relativista. No se da el caso todavía de que estemos afirmando que cualquier señal o influencia causal pudiera ser propagada con una velocidad local mayor que la de la luz.

intrínseco que esta sugerencia pueda o no pueda tener, una discusión de ella proporciona un modo instructivo de formular algunas de las cuestiones relevantes filosóficamente. Gold proporciona un ejemplo imaginario pero, sin embargo, instructivo de un sistema derivado que depende para su existencia del hecho de que el universo constituya un punto de extinción no reflexivo para la radiación [24]. Supóngase, indica, que pudiésemos colocar una estrella dentro de una caja aislada, de modo que no pudiese salir calor de la caja o entrar en ella desde fuera. Después de un tiempo suficientemente amplio se lograría el equilibrio termodinámico: volvería a la estrella exactamente la misma radiación después de su reflexión en los lados de la caja como la que sería emitida a partir de la estrella. Dentro de la caja, la flecha del tiempo se habría perdido entonces: por lo general no habría distinción entre los procesos que se diesen en la caja y los procesos inversos. Gold imagina ahora lo que ocurrirá si nosotros por un momento abrimos una pequeña ventana en un lado de la caja. Mientras abrimos la ventana saldrá radiación fuera de la caja, y, por tanto, cuando se cierre nuevamente la ventana el contenido de la caja ya no estará en un estado de equilibrio estático. Transcurrirá ahora un período de tiempo durante el cual tendrá lugar un continuo cambio de la entropía del sistema que conste del contenido de la caja (tanto la materia como la radiación). Hasta que se alcance una vez más el equilibrio, habrá una flecha del tiempo definida por la dirección en la que la entropía del contenido de la caja esté aumentando. Podemos observar que la momentánea apertura de la ventana de la caja convierte al sistema en un sistema derivado, en el sentido de Reichenbach. Supóngase, sin embargo, que el universo no fuese un punto de extinción de la radiación, de modo que cuando se abriese la ventana entrase a través de la ventana exactamente tanta radiación como saliese por ella. En este caso no se daría una disminución de la entropía. Por tanto, el sistema contenido en la caja se convierte en un sistema derivado cuando se abre la ventana debido a que el universo en su conjunto constituye un punto de extinción para la radiación. El universo constituye un punto de extinción de la radiación porque se está expandiendo. Es, por tanto, tentador suponer que si hacemos comprobaciones más profundas encontraremos que es la expansión del universo la que da razón de la "flecha del tiempo" (o, en otras palabras, la asimetría temporal que observamos al comienzo de este capítulo).

Por supuesto Gold rechaza la explicación estadística de la flecha del tiempo. Afirma: "Con toda seguridad el hecho de que tuviésemos que tratar el problema en términos estadísticos en lugar de calcular en detalle el comportamiento de todas las partes constituyentes de nuestro sistema, constituiría simplemente una falta de precisión: con toda seguridad no es rechazando información relativa a nuestro sistema como podemos hacer que nos revele el sentido del tiempo que de otro modo no mostraría" [25]. Pero, ¿es necesario que "rechacemos información" para poder aceptar la explicación estadística? Una propiedad estadística (por ejemplo

la estatura media de los estudiantes de una universidad) puede ser descrita y comprendida aun cuando poseamos conocimientos detallados de las estaturas de los estudiantes individuales. De modo semejante considérense n moléculas (que han de ser pensadas como partículas newtonianas) moviéndose al azar dentro de un cilindro provisto de un pistón. Echemos el pistón hacia atrás, doblando el volumen del cilindro. La estadística nos dirá que es casi seguro que muy pronto alrededor de $\frac{n}{2}$ moléculas estarán en el volumen original del cilindro, y alrededor de $\frac{n}{2}$ estarán en el nuevo volumen provisto por la retirada del pistón. Tal acontecimiento se muestra en la figura 16 posteriormente. La configuración *a* es seguida de la configuración *b*. Ahora inviértanse todas las velocidades de *b* de modo que se produzca *c*. *c* dará lugar a *d*, que es igual a *a* pero con las velocidades invertidas.

Ahora bien, no es cierto en absoluto que un conocimiento completo de los microestados *a*, *b*, *c* y *d* (mediante un calculador sobrehumano de Laplace) anularía nuestra apreciación de la diferencia estadística implícita en los estados de hechos mencionados. El calculador de Laplace vería que desplaziamientos arbitrariamente pequeños de las direcciones o las

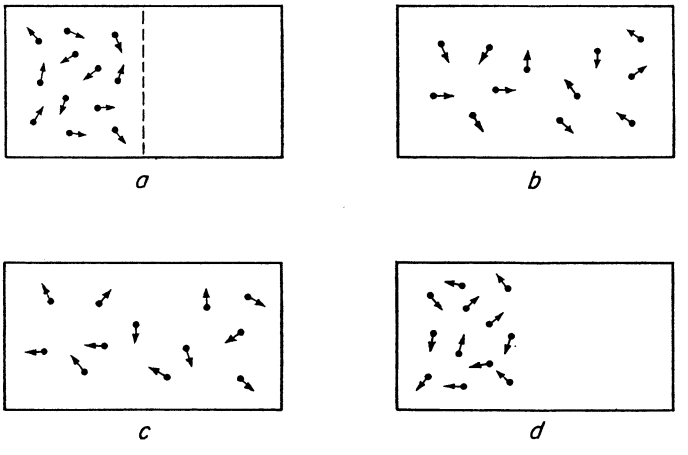

FIGURA 16.

magnitudes de las velocidades de *a* llevarían a las distribuciones uniformes de *b*, mientras que cambios arbitrariamente pequeños en las velocidades de *c* impedirían que surgiese un estado como el de *d*, en el cual todas las moléculas son segregadas en una mitad del cilindro. Por tanto, yo encuentro la objeción de Gold al tratamiento estadístico del problema de la asimetría temporal no convincente en absoluto. No es necesario que rechacemos conocimientos relativos a los detalles de un sistema para poder apreciar las propiedades estadísticas relevantes del sistema.

Ni tampoco me convence la afirmación de Gold de que el tratamiento estadístico de la asimetría temporal depende de modo importante del hecho de que el universo se esté expandiendo. No cabe duda que si estamos buscando una característica temporalmente asimétrica del universo podemos encontrarla inmediatamente en el propio hecho de la expansión. Pero, ¿en qué medida arroja esto luz acerca de las características estadísticas que dan lugar a los registros y las huellas, y, mediante ellos, a nuestra sensación subjetiva del paso del tiempo? El ejemplo de Gold muestra, en efecto, cómo en un tipo de casos la existencia de un sistema derivado depende de que el universo sea un punto de extinción de la radiación. Pero *a*) no es evidente que su explicación sea suficiente para explicar todos los sistemas derivados o, *b*) que sea necesaria tan siquiera para el tipo de casos que él imagina [15].

a) Los sistemas derivados podrían, con toda seguridad, originarse en en una gran variedad de universos posibles, algunos de los cuales serían finitos y no en expansión. Considérese el supuesto, contrario a los hechos, de un universo finito, que fue hecho en la pág. 199, anteriormente. Si suponemos que la masa-energía total de un universo finito, no en expansión, no ligado (con una geometría esférica tridimensional) es constante, entonces el firmamento en tal universo podría estar oscuro por la noche. Es cierto que la energía podría rodear al espacio, pero ninguna parte de ella podría haberlo hecho con demasiada frecuencia si consideramos que el universo no está en equilibrio termodinámico y no dista mucho de un "comienzo" en el que toda o casi toda la masa-energía estaba en forma de masa. Si no hubiese habido muestras de expansión y si el testimonio cosmológico nos hubiera llevado a postular tal universo finito no expansivo, entonces habríamos considerado que estábamos en una de las raras épocas de desequilibrio termodinámico. (Si el universo estuviese en equilibrio termodinámico no habría diferencia entre la luminosidad de las estrellas y el resto del firmamento.) Finalmente, tal universo finito no expansivo alcanzará un estado de equilibrio: suficiente masa habría sido convertida en radiación para que la densidad de la masa-energía sea la misma en todas partes. Posteriormente sobrevendrían esporádicos descensos de la curva de la entropía total del universo y finalmente el universo sufriría un proceso que sería inverso a nuestra era cósmica de ascenso. Desde luego, los seres sintientes que existiesen en esta era de descenso tendrían la impresión de que el tiempo marchaba en el sentido contrario, y, por tanto, ellos considerarían a este período de descenso y de

[15] En la discusión anterior he estado basándome en dos trabajos publicados por Gold en 1958 y 1962. Gold ha publicado recientemente otro trabajo muy asequible, "Cosmic Processes and the Nature of Time" [26], en el cual, por lo que yo puedo apreciar, presenta el mismo tipo de posición, relacionando la base estadística de la "flecha del tiempo" con la expansión del universo, aunque hay algunos aspectos nuevos. En este último trabajo, sin embargo, no utiliza el experimento imaginado de la estrella metida en una caja. En cualquier caso este trabajo me ha llegado demasiado tarde para que pudiera haberlo discutido en el presente volumen.

absorción a gran escala de la radiación, de acuerdo perfectamente con su punto de vista, como un período de ascenso y de gran emisión de radiación. Es posible, desde luego, que esta última consideración muestre la conexión entre el criterio estadístico de la "flecha del tiempo" y el principio de las ondas retardadas como fue, al efecto, utilizado por Popper. Sin embargo, es difícil ver lo que hay que decir exactamente en el caso de nuestro universo real, cuya termodinámica es más oscura que la del caso del universo finito no expansivo.

Puesto que podrían existir sistemas derivados, en el sentido de Reichenbach, incluso en nuestro universo hipotéticamente finito pero no ligado y no expansivo (y no en contracción) parecería que el sentido de la dirección del tiempo que es patente para nosotros debido al hecho de que las huellas se refieren a acontecimientos anteriores y no posteriores, y que con toda probabilidad da lugar a nuestra sensación subjetiva de la dirección del tiempo podría existir aun en ausencia del sentido de la dirección manifestada por la expansión del universo. Utilizando su ejemplo tal como él lo usa, resultaría que Gold estaría generalizando ilegítimamente a partir de un caso especial. ¿No mostraría el ejemplo de Gold de la estrella metida en la caja el mismo sentido de la dirección aun cuando el universo se estuviese *contrayendo*? En un universo en contracción la paradoja de Olbers operaría en sentido contrario. Si tuviésemos desplazamientos hacia el azul, en lugar de hacia el rojo (como sería el caso en un universo en contracción), la apertura de la ventana de la caja seguiría produciendo desequilibrio termodinámico que desaparecería gradualmente, aunque en este caso el desequilibrio sería producido por la radiación que se infiltrase en la caja, en lugar de por la radiación que saliese de ella. El progreso del sistema hacia el desequilibrio definiría exactamente la misma dirección del tiempo que en el caso del universo en expansión.

Por supuesto, presenta dificultades la noción de un universo en contracción (con desplazamiento hacia el azul). El razonamiento de la paradoja de Olbers sugiere que un universo tal (cuando menos) sería desagradablemente caluroso. Para evitar esto podríamos suponer, quizá, que la contracción no había venido teniendo lugar durante demasiado tiempo, de modo que los objetos muy distantes, que se vean muy lejanos en el pasado, no mostrarían desplazamientos hacia el azul. De todos modos, no se sigue que en un universo en contracción el "tiempo marcharía en sentido contrario" [16]. En ningún caso si "universo en contracción" se utiliza aquí en el sentido físicamente significativo de "universo que muestre desplazamientos hacia el azul". El equilibrio de un sistema puede verse tan afectado por la radiación que entra en él como por la que sale de él. Desde luego, como ya se ha señalado, en el sentido físicamente no significativo, podemos considerar a nuestro propio universo como en contracción, simplemente tomando nuestra dirección del tiempo positivo en el

[16] Vése un trabajo, "Physics and Cosmology", de H. Bondi [27]. La totalidad de este ingenioso y bellamente escrito trabajo es de interés para los intereses generales de este capítulo.

sentido opuesto de la normal. Entonces éste sería un universo en contracción mostrando desplazamientos hacia el rojo, debido a la salida de fotones *de* nuestros ojos y de las placas fotográficas, y que son absorbidos por las galaxias. Esta es únicamente otra posible descripción de nuestro propio universo.

¿Qué decir de un universo oscilante en el cual la dirección del tiempo tal como es determinada por las huellas estuviese siempre al unísono con la dirección del tiempo tal como es determinada por el criterio de la expansión? Lo que sería interesante es el demostrar que estos dos criterios podrían coincidir siempre. Es difícil determinar lo que ocurriría en un universo oscilante ya que su termodinámica es muy oscura [28].

En conjunto se diría que es la solución estadística del problema de la asimetría temporal la que es más interesante, probablemente, desde un punto de vista filosófico. Explica la asimetría temporal que es ejemplificada en la asimetría de los recuerdos y las huellas, y por ello explica también nuestra ilusión subjetiva del fluir del tiempo en una dirección.

Sugerencias respecto a lecturas posteriores

La cuestión de la asimetría temporal constituye el tema central de la obra de H. Reichenbach, *The Direction of Time* (Los Angeles: University of California Press, 1956); de la parte II de la obra de Adolf Grünbaum, *Philosophical Problems of Space and Time* (New York: Alfred A. Knof, 1963), de la obra de O. Costa de Beauregard, *Le Second Principe de la Science du Temps* (París: Editions du Seuil, 1963); y del trabajo de O. Costa de Beauregard, "Irreversibility Problems", en Y. Bar-Hillel (ed.), *Logic, Methodology and Philosophy of Science* (Amsterdam: North-Holland, 1965), págs. 313-342, con el que debe compararse el trabajo de H. Mehlberg, "Space, Time, Relativity", *ibíd.*, págs. 363-380, especialmente las páginas 373-374. Véase también el trabajo de Grünbaum, "Popper on Irreversibility" en Mario Bunge (ed.), *The Critical Approach to Science and Philosophy, In Honor of Karl Popper* (New York: Free Press, 1964), págs. 316-331. Un punto de vista completamente distinto del de los trabajos mencionados y del presente capítulo puede verse en Max Black, *Models and Metaphors* (Ithaca, N. Y.: Cornell University Press, 1962, traducción española: *Modelos y metáforas,* ed. Tecnos, Madrid, 1967), capítulo 10.

Véase también L. Boltzmann, "On Certain Questions of The Theory of Gases", *Nature,* LI (1895) 413-415; E. Schrödinger, "Irreversibility", *Proceedings of the Royal Irish Academy,* LIII (1950), 189-195; N. Wiener, *Cybernetics,* segunda edición (Cambridge: M. I. T. Press, 1961), capítulo 1; y cartas de K. R. Popper en *Nature,* LXXVII (1956), 382; LXXIX (1957), 1297, y LXXXI (1958), 402-403. Una discusión altamente matemática del problema de la irreversibilidad es la de N. G. van Kampen, "Fundamental Problems in Statistical Mechanics of Irreversible Processes", en E. G. D. Cohen (ed.), *Fundamental Problems in Statistical Mechanics* (Amsterdam: North-Holland, 1962), págs. 173-202. Una explicación elemental del teorema-H y la entropía puede verse en el capítulo 2, de la obra de Richard Schlegel, *Time and the Physical World* (East Lansing: Michigan State University Press, 1961).

Una introducción a la teoría de la información es dada en el capítulo 5 de la obra de Colin Cherry, *On Human Communication,* segunda edición (Cambridge: Massachusetts Institute of Technology Press, 1957), el capítulo 6 contiene adver-

tencias en contra de la confusión del concepto de la información en el sentido de Shannon con el concepto tal como se da en sentidos más ordinarios de la palabra, y una discusión de las explicaciones de Y. Bar-Hillel y R. Carnap de la teoría de la información semántica. Véase Y. Bar-Hillel y R. Carnap, "Semantic Information", *British Journal for the Philosophy of Science* (1953-54), 147-157; Y. Bar-Hillel, *Language and Information* (Reading, Mass.: Addison-Wesley, 1964), parte 4; y Rulon S. Wells, "A Measure of Subjective Information" en R. Jakobson (ed.), *Twelfth Symposium in Applied Mathematics* (Providence, R. I.: American Mathematical Society, 1961), págs. 237-244. Para la teoría de la información de Shannon la fuente principal es su serie de trabajos re-editados en Claude E. Shannon y Warren Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1949).

Los trabajos de T. Gold que fueron discutidos en este texto son "The Arrow of Time", en *La Structure et l'Evolution de l'Univers, Proceedings of the 11th Solvay Conference* (Brussels: R. Stoops, 1958), págs. 81-91, y "The Arrow of Time", *American Journal of Physics,* XXX (1962), 403-410. Véase también el trabajo de Gold, "Cosmic Processes and the Nature of Time", en R. G. Colodny (ed.) (Mind and Cosmos Pittsburgh: University of Pittsburgh Press, 1966), págs. 311-329.

Sobre la expansión del universo, véase H. Bondi, *Cosmology,* segunda edición (New York: Cambridge University Press, 1961). Una exposición para el lector común puede verse en H. Bondi, *The Universe at Large* (London: Heinemann, 1960), especialmente págs. 17-34.

9. DETERMINISMO, LIBRE ALBEDRIO E INTELIGENCIA

DETERMINISMO E INDETERMINISMO

En el capítulo anterior observábamos la relevancia filosófica de la mecánica estadística y la cosmología con relación al problema de "la dirección del tiempo". En este capítulo consideraremos cuestiones acerca del determinismo e indeterminismo y su relevancia respecto a la cuestión del libre albedrío.

Como definición preliminar del determinismo podemos utilizar la siguiente: Decir que el determinismo es verdadero es decir que a partir del conocimiento de las posiciones y movimientos de todas las partículas del universo en un determinado momento del tiempo, t_0, un calculador suficientemente potente podría deducir las posiciones y movimientos de todas las partículas del universo en un momento anterior o posterior. Por ejemplo, considérense los movimientos de las partículas que constituyen mi mano cuando firma un cheque de cinco dólares, como donación al club de criket de la universidad, en lugar (como podría hacer en una actitud menos meritoria) cuando realizo los movimientos de estrujar la carta del club de cricket (en la que se insinuaba lo deseable que era la donación) y arrojarla al cesto de los papeles. Según el determinismo, un conocimiento completo de las posiciones y movimientos de todas las partículas del universo permitirían a un calculador suficientemente potente deducir un conjunto de movimientos o el otro.

Esta fue, en esencia, la definición de determinismo dada por Laplace [1], y nos podemos referir al calculador suficientemente potente del parágrafo precedente como al genio de Laplace. Consideremos ahora una o dos pequeñas modificaciones de la definición.

Algunas personas podrían objetar que carece de significado el hablar de las posiciones y movimientos de todas las partículas del universo si existe un número infinito de ellas. Sea o no sea este el caso, podemos esquivar el problema de si el universo es infinito indicando que a fin de predecir un suceso en un punto P en un momento del tiempo t_0+t el genio de Laplace no necesitaría tomar en consideración acontecimientos que estuviesen a una distancia de P mayor que ct, siendo c la velocidad de la luz. Esto se debe a que la teoría de la relatividad nos asegura que ninguna influencia causal puede viajar con una velocidad mayor que c.

Otra objeción a la definición del determinismo dada anteriormente es que utiliza una proposición hipotética contra-fáctica. Si de hecho no existe ningún genio de Laplace, entonces resulta vacuamente verdadero que *sí* existe un genio de Laplace, ocurra una cosa u otra. Ocurre así en cual-

quier caso si renunciamos a las proposiciones condicionales fuertes o subjuntivas. Sin embargo, podemos evitar esta objeción señalando que el genio de Laplace es simplemente una ficción poética mediante la cual hacemos una matización puramente *lógica* acerca de una *teoría científica.* Es decir, en lugar de decir que el *mundo* es determinista digamos en primer lugar que una *teoría* es determinista. La mecánica clásica es determinista en el sentido de que los axiomas y reglas de la teoría son tales que a partir de una especificación completa de las posiciones y movimientos de todas las partículas del universo en t_0 son deducibles las posiciones y movimientos en t_0+t. La propia palabra "deducible" parece cuestionablemente modal, pero no lo es en realidad ya que si imaginamos que se formaliza la teoría, la aritmetización de la metamatemática nos permite dar una definición puramente aritmética de "deducible" [1]. Incluso en teorías no formalizadas la noción "deducible" pierde su aspecto modal si expresamos "*B* es deducible de *A*" en la forma "*B se sigue de A*".

Parece ser que podemos establecer con bastante precisión qué significa afirmar que una *teoría* sea determinista. De este modo, pocas personas querrán disputar la alegación de que la mecánica clásica es una teoría determinista mientras que la moderna mecánica cuántica es indeterminista. Puede objetarse ahora que aunque podamos establecer claramente lo que significa que una *teoría* sea determinista o indeterminista, esto no nos ayuda a establecer una tesis metafísica acerca de que el *mundo* sea determinista o indeterminista. Sin embargo, una vez que contamos con una explicación satisfactoria de lo que significa afirmar que una teoría es determinista, podemos, a partir de ahí, afirmar que significa que el mundo sea determinista. Si creemos que una teoría *T* es la verdad con relación al mundo, entonces creemos que el mundo es determinista o indeterminista según *T* sea determinista o indeterminista. Así pues, en el siglo diecinueve se creía que la mecánica clásica era verdadera, y, por tanto, se creía que el mundo era determinista, mientras que, hoy en día, creemos que la mecánica cuántica es verdadera y, por tanto, creemos que el mundo es indeterminista. Por supuesto que la mecánica cuántica puede llegar a ser sustituida por una teoría determinista, pero hasta que lo sea debemos mantener que lo más plausible es que el mundo no sea del todo determinista.

La explicación precedente del determinismo está basada en la noción de la deducibilidad formal dentro de una teoría y no hace referencia esencial a la predicción. La historia de lo que podría predecir un genio de Laplace aparece ahora como una ficción poética para establecer la cuestión formal o sintáctica. Carece, por consiguiente, de importancia para nuestra definición el que quizás en la práctica tal predicción universal no sea posible. Pudiera ser que al obtener su información relativa al estado de un sistema cualquier pronosticador real (o realizable físicamente) alteraría de tal modo el sistema que perjudicaría a sus propias predicciones.

[1] Véase la pág. 32, anteriormente.

K. R. Popper ha argumentado [2] que incluso según la física clásica no es siempre físicamente posible que un pronosticador pudiese predecir un acontecimiento dentro de un margen de error arbitrariamente pequeño. Considera este autor el caso de dos mecanismos pronosticadores *A* y *B* que se interrelacionarían en gran medida. (Estoy simplificando ligeramente la explicación de Popper, aunque confío no falsear sus características principales.)

Supóngase que *A* ha de predecir el comportamiento de *B*, y que *B* ha de predecir el comportamiento de *A*. Entonces en este caso ambos pronosticadores se interrelacionan mutuamente *en gran medida*. (Estos mecanismos pronosticadores se interrelacionan en gran medida simplemente porque un pequeño cambio en el que se predice causa un gran cambio en el que lo predice. El movimiento de un disco, por ejemplo.) Ahora bien, si *B* ha de predecir a *A*, necesita conocer su propio estado tal como está determinado por *A*. Sin embargo, si busca información relativa a su propio estado, el registro de esta información, una vez más, alterará su propio estado, y así *ad infinitum*[2]. Es decir, *B* no puede lograr nunca un conocimiento completo de su propio estado presente y, por tanto, nunca puede predecir a *A*, cuyo futuro estado también depende en parte del presente estado de *B*.

En respuesta a esto, G. F. Dear [4] ha argumentado que aunque es, por supuesto, imposible que tuviésemos dos pronosticadores *A* y *B* que se predijeran por completo mutuamente, no deja de ser cierto que dada cualquier misión predictiva finita, por ejemplo, la de predecir el comportamiento de un pronosticador *específico A*, podemos proceder a especificar un pronosticador *B* que opere con más rapidez, que sería capaz de predecir el comportamiento de *A* hasta cualquier límite preasignado de exactitud. Sea *B'* la parte de *B* que se interrelaciona con *A* (por ejemplo, brazos de medir que "sienten" los valores de los parámetros relevantes de *A*) y sea *B'* la parte calculadora de *B*. Supóngase en el momento de tiempo t_1 que *B'* está colocada en contacto con *A* y que necesita Δt_a segundos para medir los valores de los parámetros de *A*. Entonces cuando *B'* es separado de *A* y emplea Δt_b segundos en comunicar sus registros a *B''* y *B''* emplea Δt_c segundos en calcular el estado de *A* en el momento de tiempo t_2. Si $t_2 - t_1 > \Delta t_a + \Delta t_b + \Delta t_c$, *B* podrá predecir un estado de *A* posterior a su propio acto de predicción.

En cualquier caso debe insistirse en que la definición de determinismo, en términos de derivabilidad dentro de una teoría y no en términos de predicciones físicamente posibles, que ha sido dada anteriormente, parece esquivar la objeción de Popper. Dear, de hecho, define lo que significa que el mundo sea determinista, y lo hace sin referencia esencial a la noción epistemológica de las predicciones [5]. De aquí que aunque se dudara de su objeción técnica a Popper, su objeción más puramente filosófica no se

---

[2] Véase también la observación de G. Ryle en *The Concept of Mind* [3] con relación a la imposibilidad de la *auto*-predicción.

vería afectada. Es manifiesto que la caracterización del determinismo de Popper es demasiado *epistemológica.* En nuestra presente explicación podemos afirmar que el mundo es determinista o no según que creamos que la verdadera *teoría* con relación al mundo es determinista o no, y la explicación de si una teoría es determinista o no, es formal, y en el caso de teorías completamente formalizadas es incluso sintáctica, pero no es epistemológica. Aun cuando pudiera demostrarse que una teoría determinista, caracterizada formalmente o sintácticamente, implica la imposibilidad de determinadas misiones predictivas, esto por sí mismo no impugnaría el carácter determinista del mundo en cuanto definido por la teoría en cuestión.

Diremos, pues, que es razonable mantener que el universo es determinista o indeterminista según creamos o no que la verdadera teoría con relación al mundo es determinista o indeterminista. Dijimos que una teoría es determinista o no según que el estado del mundo en t_0+t sea o no deducible de una especificación completa de su estado en t_0. En realidad hay que prescindir un poco de este requisito. Por ejemplo, no consideraríamos a la mecánica newtoniana indeterminista simplemente porque el problema de la multitud de cuerpos no pueda ser resuelto analíticamente. La consideramos determinista si podemos llegar a una aproximación *de cualquier grado asignado* de exactitud del estado del mundo en t_0+t.

De acuerdo con el criterio anterior queda bastante claro que la mecánica y el electromagnetismo clásicos son deterministas, mientras que la mecánica cuántica contemporánea es indeterminista. Aquí "indeterminista" significa simplemente "no determinista según la definición anterior de determinismo". Sin embargo, existen posibles modificaciones de nuestra definición de determinismo que conservan, con todo, lo que yo considero que es el espíritu del determinismo. Por ejemplo, supóngase que el estado de todo el universo en cualquier momento de tiempo t pudiese ser deducido de una especificación de las posiciones y movimientos de todas las partículas en *dos* momento t_1 y t_2, o también, quizá, de n momentos $t_1, t_2, \ldots, t_n$. Las teorías que suscribiesen tales ideas parecerían tener bastantes derechos a ser consideradas "deterministas", aunque, desde luego, es una cuestión bastante arbitraria el que les apliquemos ese nombre o no. Un poco más alejada del determinismo de Laplace está la siguiente idea, que observé en un artículo de D. W. Sciama [6], que reconoce su deuda con K. V. Roberts. Según esta idea, las condiciones limitadoras de dos momentos t_1 y t_2 nos permiten deducir el estado de cosas no en un tiempo t cualquiera (digamos posterior a t_2) sino en un tiempo t que esté *entre* t_1 y t_2. Extendiendo de este modo la definición del determinismo, Sciama sugiere que puede conciliarse el que la mecánica cuántica sea correcta con el determinismo. Von Neumann ha demostrado que la mecánica cuántica existente resulta inconsistente si se postulan "variables ocultas" de modo que se convierta en una teoría determinista [7]. Sin embargo, Sciama afirma que en la presente sugerencia las "variables ocultas... escapan a su prohibición debido a que se refieren al futuro" [8].

La sugerencia de Sciama muestra en qué modo los científicos podrían sentirse inclinados a modificar la definición clásica de Laplace, relativa al determinismo, aun cuando nosotros pudiéramos alegar que su definición modificada se mantenía dentro del espíritu del determinismo. No es de mi competencia la discusión respecto a si la mecánica cuántica pudiera, de hecho, convertirse en determinista del modo anteriormente indicado, y, a partir de ahora, en este capítulo, consideraremos que el determinismo está definido del modo normal, mediante la referencia a la posibilidad de cálculos a partir de una especificación de las condiciones limítrofes en *un* tiempo t_0 cualquiera únicamente. Adoptemos ahora la posición más ortodoxa de que no existe ninguna extensión natural de la definición del determinismo que hará que la mecánica cuántica resulte determinista. Digamos, pues, que un acontecimiento no deducible de las condiciones limítrofes adecuadas junto con las leyes de la naturaleza ocurre por pura casualidad. "Pura casualidad" es aquí, simplemente, la negación del "determinismo".

Si el indeterminismo es verdadero, entonces algunos acontecimientos ocurren por pura casualidad. De este modo la emisión de una partícula alfa desde un átomo radioactivo en un momento de tiempo determinado es, de acuerdo con la moderna mecánica cuántica, cuestión de pura casualidad. No existe ningún estado previo del mundo que implique (*vía* las leyes de la naturaleza) que este acontecimiento ocurra en este momento. Es importante distinguir la "pura casualidad" del concepto ordinario de "casualidad" o "accidente", que es del todo compatible con el determinismo. Considérense dos acontecimientos como sigue: 1) En un momento de tiempo particular t un átomo de radio emite una partícula alfa. 2) En un momento particular de tiempo t un hombre que camina por una carretera es golpeado en la cabeza por una rama que cae. El primer caso es el de casualidad en el sentido de que, si ha de creerse a la moderna mecánica cuántica, no es deducible de ninguna especificación, por completa y precisa que sea, de ningún estado previo del universo. El segundo caso, por otra parte, es uno de casualidad en un sentido que es perfectamente compatible con el determinismo. No existe ninguna ley de la naturaleza que relacione el caminar por una carretera con el ser golpeado en la cabeza por ramas que caen, pero, sin embargo, ese suceso en particular sería deducible, si el determinismo fuese cierto, de una especificación completa de las condiciones limítrofes que se diesen en una porción suficientemente amplia del universo en algún momento previo. En términos vulgares (y menos científicos) un accidente es la intersección de dos cadenas causales. Si la esposa del hombre hubiera empleado algún tiempo más en el beso de despedida, o si la ráfaga de viento que rompió finalmente la rama hubiera venido un poco antes o después, entonces el hombre no habría sufrido el golpe en la cabeza. Esto no es incompatible con el punto de vista de que el beso de despedida de la esposa y el que se diese la ráfaga de viento estaban determinados por un estado previo del universo.

El libre albedrío y el determinismo

¿Implica la pura casualidad el libre albedrío, o la ausencia de la pura casualidad implica la ausencia del libre albedrío? Es decir, ¿son el libre albedrío y el determinismo incompatibles? Deseo mantener que la cuestión de la pura casualidad o del determinismo es irrelevante con relación a la cuestión del libre albedrío, aunque, lejos de que el libre albedrío y el determinismo sean incompatibles entre sí, se requiere una gran aproximación al determinismo en el macro-nivel para que se dé el libre albedrío.

Considérense dos computadoras *A* y *B* que realicen decisiones. Supóngase que están diseñadas para seleccionar candidatos para el cursillo de formación de personal de alguna organización. (No es mi deseo sostener aquí que la selección de candidatos debiera ser automatizada de este modo. Utilizo este ejemplo, únicamente, para aludir a una cuestión puramente metafísica.) En las entradas de las computadoras toda la información disponible es recogida mediante tarjetas taladradas. La información consta de los registros detallados de los resultados de exámenes, los grados de los caracteres mentales (por ejemplo "honrado", "moderadamente honrado", etc., "con mal carácter", "fácilmente irritable", etc.) y, en suma, toda la posible información que pueda ser de interés para un seleccionador de personas. Las máquinas han sido programadas de tal modo que quedemos satisfechos con los candidatos seleccionados (por ejemplo, ninguno de ellos es demasiado torpe, o está demasiado falto de honradez, o tiene un carácter demasiado malo).

Supóngase que existan 20 vacantes y que la computadora selecciona 19 candidatos primero. Para la vigésima (y última) plaza existen dos candidatos igualmente buenos. El último candidato es seleccionado, entonces, mediante un mecanismo de azar. Supóngase que Smith y Jones son los dos candidatos entre los cuales, de acuerdo con los criterios disponibles, la máquina no puede elegir. El mecanismo de azar es tal que proporciona una probabilidad de $\frac{1}{2}$ para la elección de cada uno de los candidatos. En el caso de la máquina *A* el mecanismo de azar consiste en una ruleta en la cual hay un contacto eléctrico. Si la rueda para en "rojo" frente a un contacto fijo eso hace que la máquina elija a Smith, mientras que si el "negro" está enfrente del contacto fijo es elegido Jones. En el caso de la máquina *B* el mecanismo de azar consiste en un contador de Geiger cercano a una pequeña cantidad de sustancia radioactiva tal que exista una probabilidad de $\frac{1}{2}$ de que el contador de Geiger sea accionado en el segundo n-avo después de empezar a funcionar la máquina, siendo n un número impar, y existe una probabilidad de $\frac{1}{2}$ de ser accionado en el minuto m-avo, siendo m un número par. Si el contador de Geiger es accionado en un segundo de número impar la máquina elige a Smith, y si es accionado en un segundo de número par la máquina elige a Jones. Supongamos que, de hecho, en el caso de ambas máquinas, Smith resulta ser el candidato elegido.

La máquina *A*, tal como fue descrita anteriormente, puede ser considerada como una máquina determinista, ya que su comportamiento puede ser descrito con bastante aproximación por medio de la física clásica. Podemos decir que la elección mediante la máquina *A* de Smith, en contra de Jones, fue determinista, ya que, con conocimientos suficientes de las entradas en la máquina y del estado inicial y funcionamiento de la máquina, habría podido ser predicha la elección de Smith en contra de Jones. Por otra parte, la máquina *B* mayúscula es una máquina indeterminista, ya que la elección de Smith dependía de un efecto indeterminista de la mecánica cuántica.

¿Fue la elección de Smith realizada por la máquina en algún grado más "racional" o menos "racional" que la misma elección hecha por la máquina *B*? Desde luego que no. ¿Fue la elección de Smith por *B* "más racional" que la elección de los diecinueve candidatos, cuyas cualidades eran tales que no fue necesario el uso del mecanismo de azar en aquellos casos? Podemos decir que la máquina indeterminista *B* no es ni más ni menos "libre" que lo es la máquina determinista *A*. Por esta razón mantengo que el principio de indeterminación de la mecánica cuántica no tiene nada que ver con el problema del libre albedrío. Sean o no sean nuestros cerebros (lo bastante cercanos a) máquinas deterministas, o contengan mecanismos de azar que actúen por "pura casualidad" no están por eso ni más ni menos poseídos de libre albedrío.

Algunos filósofos objetarían a mi presupuesto implícito anterior que la elección es cuestión de un proceso de cálculo en gran parte causal. Afirmarían que en la libre elección actuamos a partir de razones, no a partir de causas, y también afirmarían que el actuar a partir de razones ni está causado ni es una cuestión de pura casualidad. Yo considero esto ininteligible. La máquina que eligió al vigésimo hombre mediante un mecanismo de azar de la mecánica cuántica, no lo hizo debido a una causa sino que lo hizo por pura casualidad. No tenía razón alguna para elegir a un candidato u otro. Por otra parte, la máquina, al elegir los diecinueve primeros candidatos, sí se comportó de un modo determinista y podría afirmarse que las elecciones que hizo habían sido causadas. Estaba programada de tal modo que sus causas correspondían a razones, por ejemplo, estaba programada de modo que eligiese al más inteligente, sensato, etc., de los candidatos, y de este modo estaba programada para actuar de acuerdo con lo que podríamos denominar "buenas razones". Actuaba a partir de razones, precisamente por que era causada.

Puede objetarse, sin embargo, que aquí está el problema. Estaba *programada* para actuar así, mientras que nosotros actuamos libremente. En respuesta a esto podríamos hacer dos objeciones. En primer lugar, ¿no estamos nosotros mismos "programados"? Es decir, debido a nuestra herencia genética, junto con ciertas influencias ambientales, tanto cuando somos fetos como a partir de entonces, ha habido una causa que ha determinado que nuestros cerebros tuvieran una determinada estructura. Por consiguiente, calculan de determinadas maneras. Las personas que

tienen determinadas estructuras cerebrales realizarán determinadas elecciones y las personas que tengan otras estructuras cerebrales realizarán otras elecciones. La segunda objeción es como sigue: si se considera que el punto de vista de que somos, en gran medida, máquinas deterministas, implica que no tenemos libre albedrío, podemos admitir, desde luego, que tampoco tenemos libre albedrío en el sentido que el objetor tiene en cuenta. No está claro, sin embargo, cuál sea este sentido. La libre elección se supone que no es determinista ni una cuestión de pura casualidad, en el sentido en que un salto cuántico se supone que es pura casualidad. Se supone que es pura casualidad en el sentido de "no estar determinado", pero la sugerencia es de que tampoco es simplemente algo al azar y que es "puesta en acción a partir de razones". El parágrafo anterior sugiere, sin embargo, que el actuar a partir de razones no es simplemente actuar al azar, precisamente, porque es, también, actuar a partir de causas.

Es posible que el concepto de libre albedrío del hombre vulgar tenga elementos inconsistentes en él. Quizá exija del libre albedrío que sea a la vez al azar y no al azar. De ser así no nos deben doler prendas en negar la existencia del libre albedrío en el sentido del hombre vulgar. Esto no significa, sin embargo, que no tengamos a nuestra disposición los medios para hacer la mayoría de las contraposiciones que hace el hombre vulgar al afirmar o negar el libre albedrío. Un determinista puede seguir haciendo la distinción entre la condición de un hombre que va de paseo porque desea hacerlo (estando determinado por su deseo, que podemos considerar como un estado de interconexiones neurales de su cerebro) y un hombre en una celda de una prisión que no va de paseo aunque lo desea ardientemente. Este último hombre habría sido causado por el estado de su cerebro a ir de paseo pero se lo impiden los barrotes y las rejas de su celda. Ser causado por el estado interno de uno mismo en la forma primera, es, *ceteris paribus,* agradable, y el ser obstaculizado por imposiciones externas es, *ceteris paribus,* desagradable. De modo semejante, no negamos que exista una importante diferencia entre hacer una cosa *X,* porque deseamos hacerla y no tenemos deseos conflictivos, y el caso en que hacemos *X* a causa de alguna amenaza. Un hombre que da dinero a una institución de caridad porque desea ayudar a aquella buena causa actúa de un modo que es relativamente agradable. Un hombre que entrega dinero en un tribunal judicial porque tiene que pagar una multa, también, en cierto sentido, hace lo que quiere hacer, pero él desea entregar el dinero sólo en el sentido de que le desagrada todavía más la alternativa de ir a la cárcel. Una vez más el determinista puede hacer la distinción relevante, y lo haga o no mediante las palabras "libre" y "no libre", carece de importancia.

¿Qué decir de la diferencia entre un delincuente psicópata (por ejemplo un cleptómano) y un delincuente ordinario (por ejemplo un ladrón)? En ambos casos sus deseos causan su comportamiento. El hombre vulgar y la ley de lo criminal tienden ambos a contrastar el caso del cleptómano y del ladrón vulgar negando que el cleptómano fuese enteramente libre.

Podemos establecer el contraste requerido (que es todo lo que pide la ley de lo criminal) diciendo que el cleptómano no es fácilmente causado por amenazas de castigo a actuar de otro modo, mientras que el ladrón vulgar es sensible a las amenazas. A causa de esto podemos dar una justificación racional para tratar a los cleptómanos de modo distinto que a los ladrones, y podemos hacerlo sin implicar el concepto cuestionable de libre albedrío. Una vez más concluimos que si el concepto de libre albedrío del hombre vulgar es una negación tanto del determinismo como de la azarosidad, entonces su aparente inconsistencia no importa demasiado, ya que, sin embargo, podemos hacer sin su ayuda la mayoría de las contraposiciones que el hombre vulgar desea hacer.

Consideremos ahora un trozo de ciencia-ficción. En lugar de tener nuestros cerebros programados por nuestros genes y ambiente, supongamos que pudieran ser programados (o reprogramados) mediante una máquina. Supóngase que fuese posible cambiar el carácter completo de una persona aplicando un número de electrodos a su cabeza, de forma que de algún modo (de momento técnicamente inconcebible) las interconexiones entre sus neuronas (y posiblemente también todo lo que constituya su almacén de recuerdos) se altere radicalmente. Quizá pudiéramos aplicar tal máquina a un poeta humanista y convertirlo en un diabólico tecnólogo que sea un diestro inventor de armas mortíferas. Podría decirse que tal hombre "reprogramado", por poco que él estuviese manteniendo deseos contrarios, no sería libre y no merecería ni alabanza ni condena por sus actividades. En respuesta a esto, podríamos alegar que merecería tanta alabanza o condena exactamente igual que una persona normal (es decir, una persona a la que no se le hubiese cambiado el cerebro de este modo). Si fuese útil (por razones de disuasión) el castigar a una persona normal, del mismo modo sería útil (por razones de disuasión) castigar a la persona reprogramada. (Omito aquí la posibilidad de que en el caso ficticio se consiguiesen mejores resultados castigando no al programado sino al programador.) Las únicas buenas razones para castigar o recompensar, alabar o condenar, son, a mi parecer, los efectos sociales. Un filósofo que opine que es correcto condenar o castigar a una persona normal, está olvidando seguramente que la persona normal también está programada por su herencia y ambiente. Tal filósofo quizá tenga la idea peregrina, cuando piense en un asesino reincidente, que si *él* hubiese tenido el mal cerebro del criminal y su mal ambiente habría actuado, sin embargo, de distinto modo. Considero que tal posición es ininteligible.

Mi conclusión, por consiguiente, es que aunque el cerebro sea un mecanismo determinista seguimos siendo capaces de hacer las contraposiciones importantes que se señalan normalmente denominando al comportamiento "libre" o "no libre". Es una cuestión bastante arbitraria el que afirmemos o no que el concepto de libre albedrío del hombre vulgar sea contrario al determinismo, y que la posición de que la mente sea una máquina determinista (o en gran medida determinista) implique que no

tengamos libre albedrío. Mucho depende de qué hombre vulgar elijamos y de que lo cojamos o no en un estado de ánimo metafísico. Si decimos que el concepto de libre albedrío del hombre vulgar es incompatible con el determinismo, entonces, incluso en este caso, también podemos afirmar que nuestra negación del libre albedrío no implica las consecuencias alarmantes que tendría si hubiese implicado también la negación de las importantes contraposiciones que han sido discutidas anteriormente. También podemos usar "libre" y "no libre" a fin de hacer las mencionadas contraposiciones, en cuyo caso el determinismo es perfectamente compatible con el libre albedrío.

¿Qué tiene que decir la ciencia con relación a la cuestión de si el cerebro es una máquina determinista? Un buen número de filósofos ha mantenido que la naturaleza indeterminista de la mecánica cuántica abre un resquicio al libre albedrío. Nuestra discusión del libre albedrío debe haber mostrado lo que hay de erróneo en esto. Es abstractamente posible que el cerebro pudiese contener mecanismos indeterministas de disparo, como el contador de Geiger, de la máquina *B* anteriormente citada. Con todo, esto significaría únicamente que el cerebro contenía un mecanismo de azar indeterminista; si contuviese en su lugar un mecanismo de azar determinista como el de la máquina *A*, ¿en qué sentido afectaría esto a la cuestión de su libertad? No debemos descontar del todo la posibilidad de acontecimientos muy pequeños, posiblemente por debajo del nivel de la indeterminación de la mecánica cuántica que se dan en los nudos sinópticos de las neuronas simples que sean ampliados por mecanismos neuronales de modo que produzcan efectos en el comportamiento. Nada menos que un psicólogo com Sir Johm Eccles [9] ha propuesto, al efecto, tal mecanismo. Eccles piensa que de este modo los acontecimientos de una mente inmaterial pueden causar efectos en el cerebro. Posiblemente muchos científicos consideren que la idea de una mente inmaterial produce más dificultades de las que resuelve. Por lo menos la postulación de una mente inmaterial parece dar lugar a dificultades con relación a la teoría genética de la evolución mediante la selección natural. No es difícil ver, de un modo general, cómo las mutaciones de los genes llevan, *vía* la bioquímica de la embriología, a cerebros circunvolucionados de un modo ligeramente distinto y, por tanto, en último lugar, al tipo de potencialidades para el comportamiento que nos lleva a decir que un organismo tiene mente. Pero es difícil ver cómo podría reconciliarse la producción de una mente inmaterial con la explicación química de la genética. Eccles considera que la existencia de la mente inmaterial es descubierta mediante la introspección, pero aquí evitaremos la discusión de este punto que da lugar a amplias discusiones dentro de la filosofía de la mente. Observemos, sin embargo, que la postulación de una mente inmaterial no ayuda a arrojar luz sobre el problema del libre albedrío. Ya que, una vez más, podemos preguntarnos si los acontecimientos de la mente inmaterial son cau-

sados o si son al azar, o, suponiendo que no sean causados ni al azar, podemos preguntarnos en qué consiste la tercera posibilidad.

Una dificultad que creo puede originarse para quien suponga que el cerebro puede contener un mecanismo que amplifique acontecimientos muy pequeños de disparo, o quizá funcione como un todo como tal mecanismo, es la de cómo es posible que tales acontecimientos no estén enmascarados por el "ruido" [3] (en el sentido que tiene la palabra en la teoría de la información). Se dice a veces que un solo fotón puede afectar a la retina humana, y esto es sin duda verdad. No se sigue de ello, sin embargo, que la proporción de señal con ruido en el nervio óptico debe ser lo suficientemente elevada para que el funcionamiento del cerebro se vea afectado por un simple fotón que dé en la retina.

Siguiendo a D. M. MacKay [10]), por consiguiente, preferiría sugerir que el cerebro puede ser tratado como si el determinismo fuese verdadero. Incluso una simple neurona es un objeto macroscópico por lo que a la mecánica cuántica se refiere. Más aún, es posible que el comportamiento de un animal dependa del comportamiento de la masa de un gran número de neuronas, y que lo que es importante son las características estadísticas de amplios conjuntos de neuronas, y que cualquier indeterminación que pudiese mostrarse solamente en una o dos neuronas sobrantes o desaparejadas no afectaría a la cuestión.

Creatividad e inteligencia

Algunos lectores podrían hacer objeciones a la descripción de las páginas anteriores de que las máquinas *A* y *B* "hicieran elecciones" y "tuvieran razones". Los wittgensteinianos quizá deseen señalar que el "juego del lenguaje" (como ellos dirían) que jugamos con las palabras como "libre" y "elección" se juega en situaciones sociales complejas y que no podría decirse significativamente que una máquina aislada hiciese elecciones o tuviese razones. Esta objeción no parece afectar a las cuestiones metafísicas en discusión. Estén o no estén los humanos implicados en relaciones complejas entre sí, es independiente de que deban o no ser considerados como mecanismos deterministas. Las relaciones humanas complejas podrían ser relaciones complejas entre máquinas.

A veces se objeta que no podemos pensar en el cerebro humano del mismo modo en que pensamos en una máquina artificial, porque esta última "hará únicamente lo que se haya programado que haga". Puede admitirse que los seres humanos son versátiles en un sentido en que las máquinas artificiales no lo son. Un hombre puede ser un topólogo de primera categoría, un buen bebedor de cerveza y conversador, un modera-

[3] De este modo, en un sistema de transmisión eléctrica el "ruido" del canal puede deberse a movimientos casuales de los electrones (principalmente) en el circuito.

do jugador de cricket y también un no mal apreciador de la belleza femenina. Aun cuando la tecnología del futuro nos capacitase para reunir todas estas habilidades en una sola máquina, no sería muy probable que nadie desease hacerlo. Es posible que las universidades invirtiesen fondos en máquinas de topología y que los estudios cinematográficos los invirtiesen en máquinas para determinar la belleza de las damas, pero ¿quién iba a desear una combinación de ambas máquinas?

No es sorprendente que los seres humanos posean gran versatilidad. Esto es, precisamente, lo que cabría esperar de organismos que han evolucionado mediante evolución natural en vez de ser construidos artificialmente para un fin específico. Esta es la principal verdad que se esconde tras el dicho de que las máquinas harán solamente aquello para lo que han sido programadas. Nadie se ocuparía de equipar a una máquina de topología con foto-receptores, mientras que una máquina para determinar la belleza de las damas necesitaría sin duda foto-receptores con visión a todo color. Desde luego, las máquinas tienden a estar especializadas. Aun así este propósito puede ser muy general, como se muestra en el caso de computadoras versátiles, de propósitos generales. Además, excepto en ciertos presupuestos teístas, los seres humanos no han sido hechos con ninguna finalidad determinada. Ello se debe a que no han sido hechos, sino que han evolucionado. *A fortiori* no fueron hechos con ninguna finalidad.

Cuando tenemos en cuenta las anteriores preocupaciones, podemos continuar diciendo, sin demasiado temor a ser mal interpretados, que los seres humanos harán también solamente aquello para lo que han sido programados. Aquí entendemos por programas suyos: *a*) las estructuras innatas de sus cerebros y *b*) la suma total de "información" (en el sentido que tiene la palabra en la teoría de la información) que penetra en sus cerebros procedentes del ambiente, especialmente a partir de sus órganos sensoriales.

Posiblemente lo que se quiere decir cuando se afirma que las máquinas sólo harán aquello para lo que han sido programadas es que las máquinas no son capaces de pensamiento original, como lo son los humanos. Desde luego, es cierto que las computadoras existentes son más bien no inteligentes o morónicas. Con esto quiero decir que después de que han sido dispuestas para n operaciones hay una regla que especifica únicamente cuál ha de ser la operación $n+1$. Se construyen de modo que actúen de acuerdo con la regla apropiada. Compárese esto con el ejercicio del pensamiento inteligente en la solución de un problema. Considérese a un estudiante tratando de demostrar un teorema geométrico. No hay ninguna regla que le diga qué ha de hacer para buscar una prueba, aunque probablemente conoce unas cuantas estrategias que *podrían* servir. Si consigue demostrar el teorema esto se deberá a un cálculo afortunado. Se puede observar comúnmente que algunas personas son capaces de hacer cálculos afortunados de este tipo con más frecuencia que otras personas, y esto

muestra que el solucionar un problema no puede ser cuestión de simple suerte. Lo que este factor adicional pudiera ser, sin embargo, dista mucho de estar claro. Si pudiéramos contestar, en realidad, a esta cuestión, sabríamos, al menos en teoría, cómo fabricar inteligencias artificiales. Se necesita algo semejante a la suerte; un escolar que desee demostrar un teorema geométrico tiene que ponerse a buscar su solución. No es igual a hacer una división larga, donde el escolar cuenta siempre con una regla que le dice cuál es el siguiente paso, es decir, cuando existe un procedimiento de decisión. (Véanse las páginas 30-31, anteriormente.)

Supóngase que el escolar acierta con la prueba del teorema. Entonces existen reglas que le dicen que tiene la prueba. Si pudiese repasar todas las posibles argumentaciones entonces sus reglas le permitirían seleccionar una prueba del teorema en cuestión. El problema, sin embargo, radica en que el conjunto de todas las pruebas posibles (que es un subconjunto del conjunto de todas las series finitas de símbolos de su vocabulario) es infinito. Incluso si adoptásemos el presupuesto (matemáticamente injustificable): *a*) de que existe una prueba de nuestro teorema putativo y *b*) de que tiene menos de un número n de símbolos de longitud, el número de series de símbolos, a partir de las cuales tendríamos que hacer nuestra selección sería astronómico (ya que este número alcanza no a n sino a a^n, siendo n el número de símbolos del vocabulario relevante) y no sería posible ninguna selección mecánica ni siquiera realizada por la computadora más rápida. Hay que reducir el número de algún modo. Podemos descartar, por ejemplo, series de símbolos no gramaticales, series de símbolos con premisas que se haya ya demostrado que son falsas, series de símbolos que consten de pasos deductivos incorrectos, y así sucesivamente. Queda todavía sin resolver la cuestión de cómo podría lograrse en la práctica esta reducción en un tiempo bastante breve.

Si lo expresado anteriormente es correcto, entonces la inteligencia (o incluso el genio) es cuestión de *ser capaz de seleccionar*. Incluso el escribir un poema puede considerarse, de un modo muy poco romántico, como la selección de una serie de símbolos a partir del conjunto de todas las series posibles de una longitud determinada. Este es el punto de vista de W. Ross Ashby, cuyas ideas acerca de las inteligencias artificiales se basan fundamentalmente en la noción de selección [11]. Sin embargo, en un sentido (como el propio Ashby reconoce), la anterior observación resulta en sí misma bastante trivial. Shakespeare no podría haber seleccionado (pongamos por caso) un soneto de mil palabras a partir de (pongamos por caso) un vocabulario de 3.000 palabras, considerando todas las posibles combinaciones y seleccionando la mejor de entre ellas. Habría tenido que considerar $3.000^{1.000}$ posibilidades. En la proporción de una por segundo, esto le llevaría mucho tiempo, incluso en comparación con la historia de todo el universo desde la gran sacudida (suponiendo que la teoría de la gran sacudida sea cierta), cuando el átomo primitivo explotó. Sería un tiempo prohibitivamente largo, también incluso si cambiásemos el pro-

blema por el más sencillo de seleccionar cualquier soneto (no el *mejor* soneto) y si sustituyésemos a Shakespeare por una computadora que examinase las posibilidades en la proporción de una por milimicrosegundo.

Es cierto, desde luego, que nuestra cifra de $3.000^{1.000}$ posibilidades es un límite superior que podría ser reducido mediante diversos cortes. Así, supóngase que una computadora tuviese construido su interior con un diccionario con palabras de la lengua inglesa en su arsenal de recuerdos y tuviese la capacidad de comprobar si las cinco primeras letras elegidas (pongamos por caso) constituían una palabra inglesa o la primera parte de ella. Entonces si el soneto putativo empezase por "*axyba*" todas las secuencias de letras que empezasen por "*axyba*" serían rechazadas inmediatamente, sin examinarlas individualmente. Podrían añadirse nuevas limitaciones haciendo que la máquina comprobase si breves secuencias de palabras podrían formar parte de una sentencia gramatical o secuencia de sentencias, o, en casos en los que las reglas de gramática no sean afectivas, si se siguen de aquella parte de la gramática que es efectiva. (Recuérdese que decir que las reglas son efectivas es decir que existe un procedimiento de decisión para el problema de si son válidas en un caso determinado. Obsérvese que si las reglas gramaticales resultan no ser efectivas, entonces el problema de la mecanización de la gramática se convierte en un caso especial de la mecanización de la inteligencia.)

Un modo según el que podría reducirse la búsqueda exhaustiva es el de combinarlo con conjeturas inteligentes en la forma que será ilustrada en el siguiente ejemplo: supóngase que tenemos que buscar una de todas las secuencias de seis letras. Tenemos que buscar una de las 26^6 secuencias. Supóngase, sin embargo, que tenemos motivos para creer que las tres primeras letras de nuestra secuencia proceden de la primera parte del alfabeto y que las tres siguientes de la segunda parte del alfabeto. Entonces si nuestra conjetura es correcta (aunque por supuesto pudiera muy bien no serlo) tenemos 13^6 posibilidades que investigar, en lugar del número mucho mayor de 26^6. Una combinación de cálculo inteligente y búsqueda exhaustiva posiblemente resulte más potente que cada uno de estos métodos por separado. Desde luego, todavía no tenemos claro en qué consiste el cálculo inteligente. Pero el ejemplo anterior muestra cierta analogía con el modo en que el matemático pasa del problema de demostrar un teorema al de demostrar un número de lemas intermedios.

Consideremos una situación sencilla en la que se necesite de la inteligencia: el descubrimiento de pruebas de teoremas en la lógica sentencial, a partir de un determinado conjunto de axiomas. No es habitual presentar la lógica sentencial en forma axiomática, puesto que no hay necesidad de ello. Debido a la existencia de un procedimiento de decisión (el método de las tablas de verdad) podemos prescindir por completo de la necesidad de inventar pruebas. Cuando Whitehead y Russell escribieron la primera edición de *Principia Mathematica* [12], el método de las tablas de verdad no era conocido por ellos y utilizaron el método axiomático, por lo que necesitaron hacer uso de bastante inteligencia para des-

cubrir si los esquemas eran o no fórmulas válidas de la lógica sentencial.

Considérense los cuatro axiomas siguientes[4], que son suficientes para la lógica sentencial.

1) $(p \vee p) \supset p$
2) $p \supset (p \vee q)$
3) $(p \vee q) \supset (q \vee p)$
4) $(p \supset q) \supset [(r \vee p) \supset (r \vee q)]$

Un sistema axiomático también necesita reglas. Los axiomas son como las posiciones iniciales de los jugadores de ajedrez: necesitamos reglas para empezar a mover las piezas. El presente sistema requiere dos reglas: *a*) Contamos con la regla de la *separación* (detachment). Si hemos demostrado una fórmula A y también una fórmula de la forma $A \supset B$ entonces podemos escribir B como teorema. *b*) Tenemos la regla de la *sustitución*. Si A ha sido demostrada y si B resulta a partir de A mediante la sustitución de una fórmula bien formada de una letra de sentencia de A siempre que aparezca en A, entonces estamos autorizados a afirmar a B como teorema. Esta regla presupone una rigurosa definición recursiva de "fórmula bien formada", que es fácil de dar, pero que yo omitiré aquí. También necesitamos las definiciones de $A \supset B$ como $\sim A \vee B$, $A \cdot B$ como $\sim (\sim A \vee \sim B)$, y $A \equiv B$ como $(A \supset B) \cdot (B \supset A)$. Todas las posibles secuencias de sucesivas aplicaciones de estas reglas a los axiomas dan como resultado un número infinito de posibles pruebas de teoremas. ¿Cómo podemos escoger alguna prueba de algún teorema en particular? Incluso si pudiéramos presuponer un límite elevado a la longitud de la prueba, la búsqueda exhaustiva sería del todo impracticable, y, desde luego, tampoco podríamos echarnos a adivinar. Parece ser necesaria la inteligencia; quizá no excesiva inteligencia, pero alguna por lo menos. De hecho Allen Newell, J. C. Shaw y Herbert A. Simon han diseñado una máquina que posee toda esa inteligencia [16].

Supóngase que la máquina está buscando la prueba de un teorema de la forma $A \supset C$ y ya posee en su reserva de axiomas y teoremas ya probados uno de la forma $A \supset B$. Entonces la máquina podrá demostrar $A \supset C$ si puede demostrar $B \supset C$. Esto se debe a que el paso, mediante los axiomas y reglas del sistema, de las sentencias de la forma $A \supset B$ y $B \supset C$ a una de la forma $A \supset C$ es un submecanismo de la máquina. Así pues, si la máquina ha demostrado $A \supset B$ (o lo posee como axioma) y no ha conseguido demostrar directamente, mediante sustitución o separación, $A \supset C$, intentará demostrar en su lugar $B \supset C$. De modo semejante si posee $B \supset C$ pero no $A \supset B$, y si no puede demostrar $A \supset C$ directamen-

[4] Estos son los de Hilbert y Ackermann [13]. Si el segundo se cambia por "$q \supset (p \vee q)$" obtenemos cuatro de los cinco axiomas de Whitehead y Russell [14]. (El otro axioma de Whitehead y Russell fue demostrado que era redundante por Bernays [15].)

te, trata de buscar una prueba de $A \supset B$. A estos métodos se les denomina "cadenas".

Al aplicar la regla de la sustitución la máquina no irá a ninguna parte si sustituye simplemente fórmulas bien formadas tomadas al azar por letras de sentencias tomadas al azar. Tiene que aplicar sustituciones en sentencias que sean, en cierto sentido, "semejantes", a las sentencias que han de demostrarse. La máquina lleva un mecanismo incorporado que le permite calcular si dos expresiones son semejantes o no en este sentido. Esta es la *prueba de semejanza.* También existe un *proceso de emparejamiento* (matching process) que está diseñado de tal modo que, si es posible, dos sub-expresiones dadas resulten idénticas.

Puede ilustrarse, mediante el siguiente caso, cómo funcionan las estrategias mencionadas [17]. Se le ha mandado a la máquina que demuestre "$p \supset p$". Ya ha demostrado "$p \supset (p \vee p)$". Se comporta del modo siguiente: primero trata de demostrar "$p \supset p$" mediante sustitución, pero no puede encontrar ningún teorema semejante (tal como "$q \supset q$") a fin de llevar a cabo la sustitución. A continuación intenta la separación. Encuentra diversos teoremas de la forma "$(F) \supset (p \supset p)$". Por ejemplo, sustituyendo "q" por "$\sim p$" en "$p \supset (q \vee p)$" consigue "$p \supset (\sim p \vee p)$" y por tanto "$p \supset (p \supset p)$" de acuerdo con la definición (incorporada en sus reservas) de "$(F) \supset (G)$" como "$\sim (F) \vee (G)$". Sin embargo, el único axioma o teorema ya probado de la forma "$F \supset (p \supset p)$" es "$p \supset (p \supset p)$". La aplicación del método de la separación, por consiguiente, le lleva a la búsqueda de una prueba de "p". Esto, desde luego, no le lleva a ninguna parte, ya que "p" no es demostrable. Después de algunos intentos fallidos de demostrar "p" cambia entonces al método de cadenas. Busca un teorema de la forma "$p \supset (H)$" e inmediatamente selecciona el axioma "$p \supset (p \vee p)$. Esto le lleva a cambiar su problema original por el problema de demostrar "$(p \vee p) \supset p$". Comprueba los axiomas y teoremas ya demostrados y rápidamente utiliza el axioma (1) para proporcionar el resultado requerido.

Newell, Shaw y Simon, informan que en el experimento original la máquina (en realidad una computadora digital programada para simular una máquina) demostró la mayor parte de los teoremas de la lógica proposicional que habían sido demostrados en *Principia Mathemathica.* Aproximadamente la mitad de ellos fueron demostrados en menos de un minuto y la mayor parte de los restantes fueron demostrados en un espacio de cinco minutos. Unos cuantos tardaron en ser demostrados de quince a cuarenta y cinco minutos. El tiempo empleado aumentaba notablemente (quizá exponencialmente) de acuerdo con el número de pasos de la prueba. Aparecieron sorprendentes semejanzas entre el comportamiento de la máquina y el de un ser humano solucionando problemas. Por ejemplo, cuando uno de los últimos teoremas fue presentado a la máquina *después* de que hubiera almacenado en la memoria un número de teoremas que había demostrado hasta aquel momento, consiguió demostrar el nuevo teorema con mucha rapidez, pero cuando se le sacaron de la memoria estos otros teoremas no consiguió demostrar el último teorema en absoluto. De modo

semejante la máquina era suceptible a las insinuaciones; en otras palabras, el presentarle un lema apropiado intermedio, le capacitaba para continuar la demostración de un teorema que anteriormente no había conseguido demostrar.

Pudiera negarse que tal máquina poseyese ni siquiera una inteligencia rudimentaria, alegando que esa inteligencia pertenecía más bien a los diseñadores de la máquina, que idearon las estrategias de las pruebas que fueron programadas en su interior. En réplica a esto se podría preguntar: ¿no programan también padres y maestros las estrategias heurísticas dentro de los estudiantes? Sin embargo, dos muchachos de una clase de geometría pueden diferenciarse en que uno puede aplicar las estrategias y el otro no puede. Podría argumentarse, sin embargo, que incluso en este caso existe un misterio: ¿Cómo se hizo con tales estrategias el primer hombre que las aplicó? *Ex hypothesi* no contó con sugerencia alguna por parte del maestro. Debe existir alguna explicación evolucionista de cómo estrategias de pruebas tales como las que Newell, Shaw y Simon tuvieron que programar en su máquina, llegan naturalmente al cerebro humano. No hay muchas esperanzas de que pueda encontrarse una explicación plausible. Después de todo, una parte importante de las estrategias de la máquina que hemos estado considerando consiste en la tendencia a buscar y descubrir determinadas semejanzas abstractas de estructura. Esto podría basarse, posiblemente, en la habilidad de nuestro cerebro para reconocer ciertas *Gestalten,* y es del todo evidente que tal habilidad posee considerable valor para la supervivencia. Considérese la necesidad de reconocer la figura de algún animal rapaz. Cuando tenemos en cuenta la facilidad de los animales rapaces para aparecer desde ángulos inesperados e incluso para camuflarse, vemos que resulta ventajoso el que el reconocimiento de las figuras no sea una cuestión demasiado rutinaria, sino que sean asimiladas a veces figuras "casi semejantes", y a veces no. Nos encontramos en una situación bastante parecida a aquella en que modelos de pruebas semejantes son intentados y que en situaciones nuevas no resultan demasiado semejantes.

Es cierto que la habilidad para resolver problemas de la máquina de Newell, Shaw y Simon es reducida. También es cierto que la habilidad de la máquina para simular un ser humano resolviendo problemas no significa, necesariamente, que opere de acuerdo con los mismos principios que el cerebro humano. (Semejanza de funciones no implica semejanza de estructura.) En cualquier caso existe una diferencia entre la máquina y el ser humano en que la máquina trabaja más de prisa pero el cerebro humano tiene un arsenal de recuerdos más amplio. (Un problema sin resolver es el de cómo puede el ser humano extraer información tan rápidamente de su inmenso caudal de recuerdos. ¿Cómo ha de reconciliarse su amplitud con su accesibilidad?) Con todo, aunque la semejanza de funciones no implica semejanza de estructura, puede servir para indicar algo al respecto, y el conocimiento de la estructura de una máquina puede sugerir interesantes hipótesis acerca del posible modo real de operar del cerebro.

Es más sencillo suponer, a falta de posterior información, y en igualdad de circunstancias (tales como los hechos relativos a la anatomía hasta ahora conocidos), que si sólo se conoce un modo de realizar una función entonces es muy probable que el cerebro realice tal función de aquel modo.

Una importante manifestación de la inteligencia es la capacidad para aprender una lengua. Noam Chomsky ha especulado [18] que sin una tendencia innata a formar hipótesis gramaticales no podríamos jamás aprender lenguajes del tipo gramático general que Chomsky mantiene que es común a todos los lenguajes humanos. El aprendizaje del lenguaje, según su punto de vista, análogo a la formación de hipótesis en las ciencias naturales, con la excepción de que la forma general de la hipótesis está construida en nuestro interior. Mantiene este autor que sería imposible que un niño aprendiese a hablar mediante el tipo de proceso de aprendizaje inductivo sugerido por B. F. Skinner en su *Verbal Behavior* [19]. Una dificultad que nos asalta al tratar de una inteligencia artificial es el problema de cómo construir en su interior el tipo de habilidades lingüísticas que los seres humanos poseemos. Si implica incorporar un esquema "innato" (o no construido) del tipo imaginado por Chomsky entonces esto resultaría muy dificultoso, o incluso técnicamente imposible. El cerebro ofrece una gran complejidad y el imaginar el esquema requerido sería como descifrar un código difícil. Aun cuando este escepticismo en relación a los proyectos de inteligencias artificiales está justificado, no implica, desde luego, que haya nada misterioso o no mecánico en relación con el cerebro. Se trata solamente de que determinada estructura compleja ha podido ser desarrollada mediante la evolución de las especies, y resultaría difícil reproducirla artificialmente, o incluso resultaría difícil descubrir su descripción exacta. En cualquier caso estas consideraciones sugieren que la construcción de una inteligencia artificial, que sería de gran utilidad, sería una tarea de un orden de dificultad completamente distinto del de fabricar el tipo de máquina inteligente realizadora de pruebas que ha sido considerada previamente.

Con todo, una máquina tan sencilla como aquélla tiene mucho que enseñar al filósofo. Por de pronto, desmentirá a quien mantenga que una máquina sólo hará aquello para lo que ha sido programada y que no puede sorprender nunca a su diseñador. (Resulta difícil de comprender cómo tal afirmación fue considerada plausible jamás, ya que incluso los automóviles y las radios sorprenden a veces a sus diseñadores.) Una buena prueba de lo que se consideraría en un ser humano como signo de inteligencia y creatividad es la siguiente prueba de la igualdad de los dos ángulos de la base de un triángulo isósceles. Es más concisa y sencilla que la que la mayoría de los humanos conocen, y fue producida por una máquina para demostrar teoremas de geometría elemental que fue construida por H. L. Gelernter [20]. La prueba es mencionada en un trabajo de M. L. Minsky [21], y es como sigue:

Dado que en ΔABC, $AB=AC$.
Demostrar que el ángulo B=ángulo C

$$AB=AC \text{ (dado)}$$
$$AC=AB \text{ (dado)}$$
$$BC=CB \text{ (identidad)}$$
$$\therefore \Delta ABC=ACB$$
$$\therefore \text{ángulo } B=\text{ángulo } C$$

Esta prueba, en realidad, sorprende a mucha gente por su concisión. No esperan que la máquina pruebe el teorema de este modo, no habiendo pensado ellos en este método, y estando dispuestos a demostrarlo de acuerdo con el método más tosco, enseñado generalmente en los colegios, de considerar los triángulos *ABD, ACD,* siendo *D* el punto medio de *BC*. En este caso, el éxito de la máquina se debe a la propia simplicidad de su estrategia. Para mostrar que los ángulos *B* y *C* son iguales, las estrategias determinan la búsqueda de triángulos posiblemente congruentes en los que estén *B* y *C*. La máquina selecciona inmediatamente los triángulos *ABC* y *ACB*. El ser humano solucionador de problemas se confunde fácilmente en una situación tan extraña en la que los triángulos en cuestión son de hecho coincidentes. Minsky ha sugerido que cuando las máquinas son fabricadas con estrategias más potentes, también pueden confundirse del mismo modo que los seres humanos. La "superioridad" de la máquina demostradora de geometría simple en idear una elegante prueba de la igualdad de los ángulos del triángulo isósceles es, por tanto, bastante ilusoria. Sin embargo, esto no anula el hecho de que la máquina fuese capaz de hacer algo bastante bueno y sorprendente en lo cual nunca habría pensado su propio diseñador.

Ahora bien, ¿existe una diferencia únicamente de grado entre el tipo de creatividad manifestado por la máquina al idear la bella prueba de la igualdad de los dos ángulos del triángulo isósceles y el tipo de creatividad que es manifestada por el artista o el poeta al idear bellas estructuras con la pintura o con las palabras?, ¿o, acaso no existe mayor diferencia tampoco entre la actuación de tal máquina y la del científico que idea una hermosa teoría? Es prematuro contestar a esta cuestión de un modo definitivo, pero podemos ver en qué modo la fisiología, la biología y el estudio de las máquinas artificiales vienen conjuntamente a debilitar los puntos de vista de la filosofía de la mente tradicional y el dar un tratamiento mecánico a tales problemas parece ser más fructífero y más fácilmente mantenible por el progreso de la ciencia.

Se mantiene a veces que el desarrollo de una ciencia especial, a saber, la lógica matemática, muestra la imposibilidad de una filosofía mecánica de la mente. Como se observó en el capítulo 2, Gödel ha demostrado que por muchos axiomas que tengamos para la teoría elemental de los números o cualquier teoría más completa, tal como la teoría de conjuntos, a partir de los cuales pueda derivarse la teoría elemental de los números,

debe existir una sentencia verdadera de la teoría elemental de los números que no sea demostrable a partir de dichos axiomas. Algunos escritores [22] han visto en este hecho un límite a las máquinas que no atañe a las mentes de los seres humanos, porque siempre podemos convencernos a nosotros mismos de que tales fórmulas no demostrables son verdaderas.

Considérese una computadora ideal del siguiente tipo. Tiene un número finito de estados internos. Tiene una cinta magnética infinita (o si se prefiere puede estar provista de extensiones de su cinta siempre que llegue al final de ella). La cinta inicialmente contiene una secuencia finita de símbolos. Podemos considerar a los cuadrados en blanco adyacentes a los cuadrados no en blanco como conteniendo un símbolo, el símbolo cero. En cualquier momento la máquina registra un cuadrado particular sobre la cinta. Existe un conjunto finito de reglas de la forma "si la máquina está en un estado interno q_k y registra el símbolo S_i, entonces se cambia a un estado interno q_l e imprime el símbolo S_j" y "si la máquina está en un estado interno q_k y registra el símbolo S_i mueve su aparato registrador un cuadrado a la derecha [o a la izquierda]". Las máquinas ideales de este tipo se denominan "máquinas Turing". Ninguna computadora existente es una máquina Turing: las máquinas reales, por ejemplo, se gastan finalmente y, por tanto, no pueden realizar cálculos que requieran un número ilimitado de pasos. Cualquier computadora digital real, que opere bien, nunca será capaz de realizar ningún cálculo o demostrar ningún teorema que una máquina Turing correspondiente no pudiese realizar o demostrar en un tiempo ilimitado. Si construimos estrategias dentro de la máquina, como en la máquina de demostración de la lógica simple que fue comentada anteriormente, esto no permitirá que la máquina haga nada que la máquina Turing morónica no pudiera hacer dado un tiempo infinito. Así pues, se puede especificar una máquina Turing que pueda extraer todos los teoremas de un sistema formalmente axiomatizado. Por otra parte, si el sistema no tiene procedimientos de decisión, es matemáticamente imposible que exista una máquina que extraiga todos los no teoremas, porque, si existiese tal máquina, podríamos contar con un procedimiento de decisión para saber si una sentencia era un teorema o no: esperando simplemente a que o bien la máquina productora de teoremas o la máquina no productora de teoremas extrajese la sentencia. Cuando no existe el procedimiento de decisión, podemos especificar únicamente una máquina extractora de teoremas. Supóngase que después de unos cuantos millones de pasos todavía no ha sacado la sentencia en cuestión. Evidentemente no sabemos si esto se debe a que nunca lo hará o simplemente a que no ha estado operando el tiempo suficiente. Ahora bien, una máquina con estrategias incorporadas para pruebas dentro del sistema formal quizá producirá simplemente debido a una racha de suerte un teorema que la máquina morónica produciría *finalmente*. En el caso de que la máquina morónica no produjese jamás el teorema, tampoco lo haría la máquina con extrategias incorporadas.

Considérese una máquina Turing que extraiga teoremas de un sistema

formal de la teoría elemental de los números (o alguna teoría más completa). El teorema de Gödel nos asegura que si el sistema correspondiente a la máquina Turing es consistente, entonces existirá una sentencia que no sea demostrable ni no demostrable por la máquina. Ahora bien, los antimecanicistas argumentan del modo siguiente. Sea cual sea el sistema y sea cual sea la máquina, nosotros, en cuanto seres humanos, podemos demostrar que G es en realidad verdadera. Por tanto, podemos estar seguros de que G sea verdadera, aun cuando G sea una proposición indecidible del sistema correspondiente a la máquina Turing que me representa, en el sentido de ser capaz de demostrar todos los teoremas que yo puedo demostrar, y, por consiguiente, parecería que esta máquina Turing sería a la vez capaz y no capaz de demostrar G. Esto lleva, mediante *reductio-ad-absurdum* a la proposición de que no puede existir ninguna máquina Turing que me represente.

El fallo de la argumentación, como ha sido indicado por Hilary Putnam [23], radica en el hecho de que no sabemos que G sea indemostrable. Sólo sabemos que G es indemostrable si la máquina Turing que me representa es consistente. Ya que lo que Gödel demostró no fue un teorema de la forma:

$$G \text{ es indemostrable}$$

sino uno de la forma

$$\text{consis} \supset G \text{ es indemostrable}$$

en la que "consis" es la sentencia metamatemática que afirma la consistencia del sistema en cuestión. A menos que podamos demostrar la consistencia [5] no podemos demostrar la indemostrabilidad de G. Ahora bien, ¿es probable que yo pueda demostrar la consistencia de una máquina tan complicada que pudiera sustituirme? (Aun cuando fuesen permitidos métodos transfinitos [6].) Y si no se permiten los métodos transfinitos es matemáticamente imposible demostrar la consistencia por lo menos de un sistema tan completo como el de la teoría elemental de los números. (Esto es un corolario del propio teorema de Gödel.) Desde luego si el sistema es *inconsistente,* entonces G es ciertamente demostrable, ya que cualquier cosa (incluidas a la vez G y $\sim G$) puede ser derivada de una contradicción. Véanse las págs. 32-33.

Desde luego es, con toda seguridad, del todo implausible que se pudiera demostrar con detalle nada en relación con una máquina Turing que me representase, en el sentido de ser capaz de demostrar todos los teoremas que yo podría demostrar durante toda mi vida, a menos, por supuesto, que tomemos en cuenta los razonamientos inconsistentes que he hecho

[5] En realidad, como se recordará, Gödel utilizó en su trabajo original el requisito más completo de la consistencia ω, pero Rosser demostró que el requisito más incompleto de la consistencia sirve.

[6] Como en la prueba de Gentzen de la consistencia de la teoría elemental de los números [24].

durante mi vida, en cuyo caso cualquier máquina inconsistente servirá trivialmente. Con toda seguridad la argumentación adquiere su plausibilidad a causa de que consideramos únicamente el caso de una máquina Turing que me representa en un instante de mi vida, y luego olvidamos que es casi seguro que esta máquina Turing no me representará en un instante posterior, debido a la continua corriente de "información" que penetra en mi cerebro. Desde luego, la verdad captada intuitivamente de una sentencia de Gödel corresponde a una argumentación que de formalizarse tendría que hacerse en un metalenguaje semántico. Ahora bien, en tanto que podamos hacer esto, parecerá que nuestro conocimiento intuitivo de las reglas sintácticas y semánticas del lenguaje que hablamos procede de nuestra habilidad para formular hipótesis empíricas basándonos en nuestras experiencias del lenguaje que hablamos. Basándonos en estas experiencias podemos, con toda seguridad, ascender también a la semántica de nuestro metalenguaje semántico, y así *ad infinitum*. De este modo podemos correspondernos, instante por instante, con máquinas Turing cada vez más potentes. No hay nada que no sea mecánico en esto, sino únicamente el hecho trivial de que estamos siendo "re-programados" continuamente, como resultado de las experiencias de nuestro propio comportamiento, así como de todas las demás cosas. Obsérvese que a fin de demostrar cualquier cosa con relación a una máquina Turing que me represente en un instante t tendríamos que descubrir cuál era ésta, probablemente descubriendo los detalles de mi circuito neuronal en t. Aun cuando pudiera hacerse realidad este trozo de ciencia-ficción, seguiría siendo evidente que el propio acto de determinar este circuito lo cambiaría, y, por consiguiente, una vez que conociéramos qué máquina Turing me representaba en t, yo ya no sería representado por ella. No hay nada dentro de la teoría mecánica de la mente que sea en lo más mínimo incompatible con este hecho.

Mecanismos intencionales (purposive mechanisms)

El mecanismo realizador de pruebas que fue considerado en la sección precedente proporciona una muestra particular de un mecanismo intencional (purposive) o "buscador de un fin". Los mecanismos intencionales pueden ofrecer diversos grados de complejidad, y presentar diversos grados de sutileza en la consecución de sus fines. En el nivel más bajo tenemos mecanismos tan simples como un termostato o el regulador de una máquina de vapor. En estos casos la diferencia entre el estado real de los hechos y el objetivo (entre la temperatura existente y la temperatura requerida) produce un cambio en el mecanismo que causa un cambio en el estado real de los hechos en el sentido de acercarlos más al objetivo. Normalmente esta acción compensadora irá más allá de lo marcado, produciendo un cambio en la dirección opuesta, de modo que lo que realmente obtenemos es una temperatura o un índice de revoluciones que se separan

ligeramente de los requeridos. No son mucho más complejos los mecanismos de los tropismos, tales como el mecanismo que hace que una polilla vuele hacia la luz, o el que hace que una "tortuga" mecánica de Grey Walter busque su re-cargador de baterías [25]. Tenemos después los corredores mecánicos dentro de un laberinto que intentan al azar tipos alternativos de acción hasta que (tal vez) consiguen salir del laberinto, y, en un nivel superior de elaboración, corredores de laberinto que aprenden de la experiencia, en el sentido de que ya no intentan tipos de acción que no hayan tenido éxito. Considérese también la "tortuga" de Grey Walter modificada con su mecanismo de reflejos condicionados electrónico auxiliar. Después tenemos también mecanismos con computadoras incorporadas, ejemplificados por el cañón antiaéreo predictor controlado. El predictor recibe información mediante radar del curso del avión que sea blanco para disparar, y calcula la posición en que estará el avión en el momento en que la bomba llegue hasta él, y hace que el cañón dispare a esta posición calculada. Desde luego, en una era de misiles automáticamente controlados y de satélites espaciales, así como de fábricas y refinerías de petróleo automáticas, está ya resultando del dominio público el que las máquinas pueden realizar muchas de las actividades intencionales de los hombres, y que incluso pueden realizarlas mejor. Tales hechos han contribuido, al menos tanto como cualquier análisis puramente filosófico, a debilitar la posición que mantenía que en teleología hay algo misterioso o incompatible con las ideas mecánicas.

Podría preguntarse si la diferencia entre mecanismos intencionales y no intencionales es muy grande. Richar Taylor ha considerado recientemente [26] los casos de: 1) un torpedo que es guiado por las ondas sonoras de la hélice del barco objetivo de modo que se "instale" en el objetivo, y 2) una bomba, conectada con un barco mediante una cuerda, de modo que cuando el barco se detiene la bomba cae en el barco. Si denominamos al primer mecanismo intencional, ¿por qué no al segundo? Además, ¿no sugiere esto que el adscribir intencionalidad al torpedo automático es tan erróneo como lo sería la creencia de una persona que no observase la cuerda de la bomba y pensara que la bomba estaba buscando intencionalmente al barco objetivo? En contra de Taylor yo sugeriría que existen importantes diferencias entre estos dos casos. Podríamos, tal vez, considerar al barco y la bomba *juntos* como a un mecanismo intencional extremadamente simple, pero la propia bomba no tiene por qué contar con ningún mecanismo en absoluto. Lo que se transmite del barco al torpedo es información en forma de ondas sonoras, pero estas ondas sonoras no producen ellas mismas la fuerza que mueve al torpedo hacia la posición predicha del barco. En el caso de la bomba, sin embargo, la cuerda que podría, tal vez, considerarse que transmite "información" con relación al barco es, en realidad, la misma cosa que mueve la bomba. El sistema del barco y la bomba es un sistema de "flujo de energía" más que un sistema de "flujo de información" [27]. Es esta característica, con toda seguridad,

la que haría a los cibernéticos considerar al torpedo, y no a la bomba, como un mecanismo intencional.

Los movimientos del barco y de la bomba a la que está unido pueden considerarse "epistemológicamente dependientes" entre sí, ya que existe una ley simple que relaciona los movimientos del barco y de la bomba. Esta noción de "dependencia epistemológica" (epistemic dependence) se debe a G. Sommerhoff [28], que la utiliza para analizar la diferencia entre casos de aparente intencionalidad debida a la estabilidad física, como aquélla en que una pelota regresa siempre al centro de una concavidad hemisférica, y los casos de dirección de un objetivo. Sommerhoff admite que la noción de dependencia e independencia epistemológica no es absoluta, y, por tanto, quizá también deba admitirse que no existe una diferencia tajante y absoluta entre los mecanismos intencionales y los no intencionales. Sin embargo, yo no comprendo por qué los oponentes y defensores de la noción de que algunos mecanismos puedan ser intencionales tienen que mantener que la diferencia es tajante. Quizá no sea más tajante que la diferencia entre lo vivo y lo no vivo. ¿Es un virus materia viva o no viva? Las dudas relativas a denominar o no denominar a un virus "vivo" no nos harán dudar en denominar a un gato objeto vivo. De modo semejante, las dudas respecto a denominar a la colocación de la bomba un mecanismo intencional no nos tienen por qué llevar, necesariamente, a dudar de que el torpedo autoguiado lo sea.

También puede resultar instructiva otra de las objeciones de Taylor a la noción de mecanismo intencional. El pone el ejemplo de una anciana dedicada a enhebrar agujas en una producción en cadena. Ella es muy hábil y mete el hilo en el agujero de la aguja la mayoría de las veces. Posteriormente, debido a rencillas con su jefe, deja de enhebrar las agujas con mucha frecuencia. Supóngase ahora que es sustituida por un sistema automático de enhebrar agujas. ¿Cómo podría fabricar un ingeniero un enhebrador automático de agujas que no enhebrase el hilo a menudo *a propósito*? Evidentemente no sería buena cosa fabricar una máquina que no enhebrase el hilo frecuentemente. Eso sería únicamente, desde luego, fabricar un mecanismo de enhebrar agujas ineficaz.

Con toda seguridad esto no demuestra que la anciana no sea un mecanismo intencional. Muestra, simplemente, que a causa de la herencia y el ambiente, ha llegado a tener muchos fines, o "programas" realizadores de objetivos. Es, por ejemplo, no sólo un mecanismo enhebrador de agujas, sino también un mecanismo para vengar rencillas. (Si las circunstancias hubieran sido diferentes, y ella hubiera sido la encargada de prepararle el té a su jefe, esta última circunstancia le habría llevado a vengar sus rencillas de modo diferente, por ejemplo sirviéndole el té frío a su jefe.) Puesto que las rencillas pueden ser vengadas de tantos modos diversos, un mecanismo para vengar rencillas, semejante al humano, sería mucho más difícil de confeccionar que un mecanismo para enhebrar agujas.

Es cierto que en el caso de los seres humanos, por lo menos, parece ser que nos inventamos nosotros nuestros propios objetivos. Esto nos

lleva de nuevo a la cuestión general de la creatividad y la inteligencia. Con todo, la invención de un nuevo objetivo no es, necesariamente, nada misterioso. Existe un sentido en el cual incluso un mecanismo burdo del todo puede proponerse a sí mismo un objetivo no designado por su creador, como cuando el mecanismo de posición de un cohete espacial se pasa a la estrella luminosa que no debe. También debe recordarse que no dejamos de hablar de intencionalidad en los humanos cuando *no* inventan sus propios objetivos. Existe algo profundamente no original en la mayoría de los objetivos humanos. Considérese el común deseo de casarse y tener hijos. Algunas personas, desgraciadamente, carecen de los genes que las harían susceptibles a las presiones sociales en este sentido, pero la mayoría de las personas las tienen y, con toda seguridad, podemos considerar que los objetivos de tales personas son programados en ellas tanto genética como ambientalmente. También debe preguntarse si la apariencia de objetivos más originales no se debe simplemente a medios originales de resolver el problema de conseguir objetivos más generales y carentes de originalidad. De este modo podría ser programado un deseo de crear un nuevo tipo de matemáticas dentro de un inteligente y potencialmente ambicioso matemático mediante amantes padres, maestros animosos, amigos ya ambiciosos y demás. El objetivo de crear *este* tipo de matemáticas es evidentemente nuevo, pero no lo es necesariamente el objetivo de crear una nueva y bella teoría.

Por muy lejos que nos encontremos de una comprensión de las formas más bien sencillas de solucionar problemas, nos encontramos todavía más lejos de la comprensión de las formas más elevadas de la creatividad en la ciencia, la literatura y demás. Con todo, intentos posteriores de teorizar mecánicamente en relación con estas actividades tenderán, con toda probabilidad, a verificar cambios en el clima de opinión, con lo que llegaremos a no estar de acuerdo tan siquiera con que el genio sea algo misterioso. Después de todo, incluso los fallos de las teorías mecánicas de la mente son fallos airosos. El creador de misterios no se expone a posibles fallos, y, por consiguiente, el éxito verdadero no es tampoco posible para él. Las teorías mecánicas de la mente también nos dan una idea de a qué se parece una explicación real. El recurrir a ideas vitalistas o espiritualistas resulta menos atractivo si nos damos cuenta que los defensores de tales ideas tienen la obligación de mostrarnos con detalle cómo pueden ayudarnos sus ideas a explicar el comportamiento humano. No deben esperar contentarnos simplemente con la vaga sugerencia de que, después de todo, un espíritu puede hacer *cualquier cosa*[7]. No sirve de mucho el sustituir los problemas no solucionados relativos al funcionamiento de mecanismos por problemas no solucionados relativos al funcionamiento de espíritus. Por supuesto, al igual que una máquina Turing podría ser realizada aproximadamente por un ángel que operase de acuerdo con re-

[7] Después de todo, si un espíritu puede atravesar puertas cerradas, es de esperar que pueda realizar cualquier otra cosa que consideremos misteriosa, tal como escribir poemas.

glas rigurosas. De modo semejante incluso los problemas de lo espiritual podrían convertirse en problemas cibernéticos [8].

Al determinar la probabilidad de las explicaciones mecanistas futuras de las habilidades humanas es necesario tener en cuenta la posibilidad de un determinado tipo de falacia. Supóngase que hemos recorrido el 80 por 100 [9] del camino hacia la realización de algún cometido (supongamos la traducción completamente mecánica y automática del ruso al inglés). Y. Bar-Hillel, en una crítica instructiva de la posibilidad de la traducción a máquina de alta calidad, totalmente automática [30], menciona la "falacia del 80 por 100", descrita por W. E. Bull, C. Africa y D. Teichrow [31], de suponer que se sigue, por consiguiente, que un cuarto del esfuerzo ya realizado nos hará recorrer el resto del camino. Así pues, en el caso de la traducción por máquinas, ya es posible mostrar máquinas que traducen de tal modo que un director de un periódico, inteligente, puede hacer una buena traducción a partir de la traducción hecha por la máquina. Sin embargo, para hacer la posible traducción a gran velocidad, totalmente automática, se originan dificultades de un orden distinto de aquellas que hasta ahora han sido superadas. Bar-Hillel da una explicación muy sencilla de esto. Supóngase que una computadora electrónica está traduciendo la sentencia inglesa "the box is in the pen" en un lenguaje que tiene distintos significados para "pen", a saber, "utensilio para escribir" y "cercado". Un traductor humano no dudaría en traducir la anterior sentencia por "la caja está en el cercado", y no "la caja está en la pluma", debido a su *conocimiento* de que las cajas no suelen ser lo suficientemente pequeñas para caber dentro de plumas estilográficas, sino que se encuentran generalmente dentro de patios o cercados. Para hacer factible tal traducción de elevada calidad habría que construir no solamente las reglas gramaticales y semánticas de los lenguajes dentro de la computadora sino también unos conocimientos enciclopédicos, y habría que resolver también el problema de recoger información rápidamente a partir de tan amplias reservas. (Este último problema, desde luego, constituye uno de los problemas sin solucionar que resultan un impedimento para la explicación mecánica de la inteligencia humana.)

Ahora bien, en correspondencia a esta "falacia del 80 por 100" dentro de la tecnología existe la posibilidad de una "falacia del 10 por 100" dentro de la metafísica. Debido a que el pensamiento cibernético ha proporcionado explicación de algunas actividades inteligentes e intencionales de los seres humanos, es natural saltar de esto a la conclusión de que puede explicarlas todas. Con todo, uno quizá se sienta menos pesaroso de cometer la "falacia del 10 por 100". Ya que, mientras que en tecnología se pueden ofrecer buenas razones científicas de por qué el último 20

[8] Un divertido ejemplo que ilustra cómo las ideas básicas de la cibernética podrían ser aplicadas a lo espiritual puede verse en la obra de W. Ross Ashby, *An Introduction to Cybernetics* [29], 4/15.

[9] Desde luego, un porcentaje preciso tal como el "80 por 100", carece probablemente de significado en un contexto como éste, pero esta cuestión carece de importancia aquí.

por 100, o incluso el 2 por 100, de un proyecto pueden originar dificultades de un nuevo orden de magnitud, y, por consiguiente, el 100 por 100 total de la tarea no puede ser llevada a cabo, la falacia del 10 por 100 en metafísica sería objetable sólo si se pudiese demostrar que existía una alternativa a los métodos mecánicos de explicación que se encontrase en mejor situación que éstos. Puesto que no se ha presentado todavía ninguna alternativa científicamente respetable a las teorías mecánicas de la mente, no es del todo irrazonable continuar manteniendo el presupuesto de que nuestro índice de éxito del 10 por 100 se debe únicamente a que no somos lo suficientemente inteligentes o a que todavía no poseemos suficientes conocimientos.

La psicolingüística y el problema kantiano

La lingüística moderna quizá sólo sea aventajada por la filosofía en sus ramificaciones interdisciplinarias, y no es una simple coincidencia el que algunos de sus máximos exponentes, principalmente Chomsky, sean filósofos muy capacitados. Considérese el papel que desempeña la noción de recursividad (recursiveness) dentro de la obra de Chomsky. Su *Sytantic Structures* [32] presenta evidente vinculación con la lógica matemática. Los problemas de la traducción mediante máquinas unen el tema con el de la electrónica y con la teoría de las computadoras. La lingüística también se vincula a la psicología, ya que da lugar a cuestiones relativas a cómo los niños aprenden la gramática de una lengua, y a si esto puede ser o no ser explicado sobre la base de la teoría ortodoxa del aprendizaje. También está vinculada, por supuesto, con las disciplinas humanísticas filosóficas. Es una gran disciplina, por tanto, para quien desee acabar con los departamentos desvinculados de la universidad, y quizá es la que pudiera ayudar a derrocar lo que C. P. Snow [33] ha denominado "Las dos culturas" [10].

Ahora bien, el impacto de la lingüística en la psicología es de particular interés filosófico debido a la sugerencia de Chomsky, a la que nos hemos referido en la pág. 237, anteriormente, de que pudiera haber un esquema innato en la estructura de nuestro cerebro que lleve al niño a formular hipótesis de un tipo particular de gramática. Esto suscita la cuestión semejante de la posibilidad de que los esquemas conceptuales sean innatos; por supuesto, las dos cuestiones no son del todo separables, ya que la gramática de nuestro lenguaje refleja un esquema conceptual. El modo en que un lenguaje natural tiende a utilizar sustantivos y adjetivos, en lugar de cuantificadores, variables y predicados, como en un lenguaje artificial basado en la lógica matemática, nos inclina, evidentemente, a una metafísica de sustancias y cualidades cambiantes [11]. Imma-

[10] Es decir, las ciencias por una parte y las humanidades por la otra.

[11] P. F. Strawson [34] ha argumentado con bases filosóficas a favor de la inevitabilidad de categorías metafísicas basadas en las categorías vulgares.

nuel Kant mantenía que ciertas categorías, como la de sustancia y la de causa, eran inevitables. La ciencia moderna ha demostrado, en realidad, que Kant se equivocaba a este respecto, al prescindir del uso de estas categorías en teorías fundamentales, pero quizá pudiera argumentarse que tenemos que utilizar el lenguaje vulgar como un paso previo para un lenguaje científico más adecuado. Es, desde luego, difícil imaginar cómo pudiera un niño empezar a aprender un lenguaje que contuviese la teoría de la cuantificación y el concepto de los objetos como sólidos espacio-temporales cuatridimensionales, de tal modo que el niño hablase de pronto de las diferencias entre secciones transversales de conejos, en lugar de conejos cambiando. Ahora bien, la cuestión de en qué grado dependen nuestros esquemas ordinarios de la estructura innata del cerebro, es lo que yo denomino el problema kantiano; y, a mi parecer, tal problema será resuelto finalmente por la neurofisiología y por la psicolingüística. Por supuesto, aunque se demostrase que nuestro esquema conceptual ordinario de sustancia y cambio es innato en nosotros, esto no significa que sea necesariamente bueno. Significa que ha resultado útil en la historia del género humano, no que sea metafísicamente adecuado. Observábamos en el capítulo 3 que una teoría incorrecta puede, a veces, ser más útil en la práctica que una que sea correcta.

Sugerencias respecto a lecturas posteriores

La teoría liberal del libre albedrío es quizás defendida mejor que por nadie por C. A. Campbell en su "Is Free-Will a Pseudo-Problem?", *Mind,* LX (1951), 441-465. Posteriores consideraciones sobre el libre albedrío, desde el punto de vista de su plausibilidad científica, son ofrecidas por J. J. C. Smart, "Philosophy and Scientific Plausibility", en Paul K. Feyerabend y Grover Maxwell (eds.), *Mind, Matter and Method, Essays in Philosophy and Science in Honor of Herbert Feigl* (Minneapolis: University of Minnesota Press, 1966), págs. 377-390. Muchos filósofos niegan que exista ninguna incompatibilidad entre el libre albedrío y el determinismo. Un bello artículo de R. E. Hobart (Dickinson S. Miller) es "Free-Will as involving Determinism and Inconceivable without It", *Mind,* XLIII (1934), 1-26.

Sobre cuestiones relacionadas con la mecanización de la inteligencia, existen diversos útiles artículos en el National Physical Laboratory Symposium, *Mechanisation of Thought Processes* (London: Her Majesty's Stationery Office, 1959), y C. E. Shannon y J. McCarthy (eds.), *Automata Studies* (Princeton: Princeton University Press, 1956). Véase también el artículo "Artificial Intelligence", de Marvin L. Minsky, *Scientific American,* CCXV, 3 (septiembre, 1966), 246-260. Dos trabajos muy útiles de Herbert A. Simon son "Thinking by Computers", en R. G. Colodny (ed.), *Mind and Cosmos* (Pittsburgh: University of Pittsburgh Press, 1966), págs. 3-21, y "Scientific Discovery and the Psychology of Problem Solving", *ibíd.,* págs. 22-40. F. H. George, *The Brain as a Computer* (New York: Pergamon Press, 1961), es un útil tratado. Una comparación entre las computadoras electrónicas y los cerebros humanos puede verse en J. von Neumann, *The Computer and the Brain* (New Haven, Conn.: Yale University Press, 1958).

Sobre los mecanismos intencionales véase W. Ross Ashby, *Design for a Brain,* segunda edición (New York: Wiley, 1960, y en rústica, New York: Barnes and Noble, 1960, traducción española: *Proyecto para un cerebro,* ed. Tecnos, Madrid,

1965), e *Introduction to Cybernetics* (New York: Wiley, 1961, y en rústica, 1963), G. Sommerhoff, *Analyticial Biology* (London: Oxford University Press, 1950) y la discusión de los encadenamientos teleológicos y causales en R. B. Braithwaite, *Scientific Explanation* (London: Cambridge University Press, 1953, traducción española: *La explicación científica,* ed. Tecnos, Madrid, 1964). Véase también J. J. C. Smart, "Causality and Human Behaviour", *Aristotelian Society Supplementary Volume,* XXXVIII (1964), 143-148, que es parte de un simposio con D. W. Hamlyn. Argumentos en contra de la posición mecanista pueden verse en Richard Taylor, *Action and Purpose* (Englewood Cliffs, N. J.: Prentice-Hall, 1966), capítulo 15. A. R. Anderson (ed.), *Minds and Machines* (Englewood Cliffs, N. J.: Prentice-Hall, 1964) incluye la reedición de una serie de artículos, incluyendo el famoso artículo "Computing Machinery and Intelligence", de A. M. Turing. J. von Neumann, "The General and Logical Theory of Automata", en L. A. Jeffress (ed.), *Cerebral Mechanisms in Behavior, The Hixon Symposium* (New York: Wiley, 1951), págs. 1-41, contiene una discusión de un autómata capaz de autorreproducirse. P. Benacerraf, "God, the Devil and Gödel", *Monist,* LI (1967), 9-32, es una excelente e ingeniosa discusión de las argumentaciones que tratan de utilizar el teorema de Gödel para refutar el mecanismo. Un trabajo interesante de Theodore Mischel, "Pragmatic Aspects of Explanation", *Philosophy of Science,* XXXIII (1966), 40-60, es de interés en cuanto relaciona una discusión de la teoría de la explicación científica de Feyerabend (véase el capítulo 3 anterior) con una discusión de la relación entre las explicaciones mecanistas y las explicaciones vulgares en términos de objetivos.

La actitud de N. Chomsky respecto a la psicolingüística aparece en su revisión de la obra de Skinner, *Verbal Behavior in Language,* XXXV (1959), 26-58. Esta revisión ha sido reeditada en J. A. Fodor y J. J. Katz (eds.), *The Structure of Language, Readings in the Philosophy of Language* (Englewood Cliffs, N. J.: Prentice-Hall, 1964), págs. 547-548. Véase también Noam Chomsky, "Recent Contributions to the Theory of Innate Ideas", *Synthèse,* XVII (1967), 2-11, Hilary Putnam, "The 'Innateness' Hypothesis and Explanatory Models in Linguistics", *ibíd.,* 12-22, y Nelson Goodman, "The Epistemological Argument", *ibíd.,* 23-28. Sobre el "problema kantiano" véase la observación de la pág. 168 de J. J. C. Smart, "Conflicting Views about Explanation", en R. S. Cohen y M. W. Wartofsky (eds.), *Boston Studies in the Philosophy of Science,* Vol. II (New York: Humanities Press, 1965), y págs. 234-235 del trabajo de P. K. Feyerabend, "Reply to Criticism", en el mismo volumen. Aquí Feyerabend llama la atención sobre el parágrafo 58 (b) del trabajo de Wilfrid Sellars, "Scientific Realism or Irenic Instrumentalism", *ibíd.,* 171-204.

NOTAS REFERENCIALES

Capítulo 1:

1. Véase el trabajo de J. Holland Rose, "Was the Failure of the Spanish Armada due to Storms?", *Proceedings of the British Academy,* XXII (1936), 207-244, especialmente en la pág. 226. (Hay una errata de imprenta en la segunda nota, al pie: en lugar de "E. M. Smart", debería ser "W. M. Smart".)
2. Israel Scheffler, *The Anatomy of Inquiry* (New York: Knopf, 1963), pág. 5.
3. *Ibíd.,* pág. 6.
4. Véase, por ejemplo, Noam Chomsky, *Syntactic Structures* (The Hague: Mouton, 1957).
5. W. H. Watson, *On Understandign Science* (London: Cambridge University Press, 1938).
6. *Ibíd.,* pág. 8.
7. Gilbert Ryle, *Dilemmas* (London: Cambridge University Press, 1954).
8. L. Wittgenstein, *Tractatus Logico-Philosophicus,* traducción de D. F. Pears y B. F. McGuinness (London: Routledge & Kegan Paul, 1962).
9. W. V. Quine, *Word and Object* (Cambridge: M. I. T. Press, 1960).
10. *Encyclopedia Britannica,* 13.ª edición, XXI (1926), pág. 441.
11. Véase H. Feigl, "De principiis non Disputandum...?", en Max Black (ed.), *Philosophical Analysis* (Ithaca: Cornell University Press, 1950).
12. Albert Einstein, "Autobiographical Notes", en P. A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist,* 2.ª ed. (La Salle, Ill.: Open Court, 1951), especialmente pág. 53.

Capítulo 2:

1. Una explicación del método de Gödel para demostrar la completitud puede verse en D. Hilbert y W. Ackermann, *Principles of Mathematical Logic,* traducido por L. M. Hammond, G. G. Leckie y F. Steinhardt y editado con notas por R. E. Luce (New York: Chelsea, 1950), págs. 92-101. Un método mucho más sencillo para demostrar la completitud de la lógica de predicados aparece en la obra de W. V. Quine, *Methods of Logic,* edición revisada (New York: Holt, Rinehart & Winston, 1959), págs. 253-260.
2. Alonzo Church, "A Note on the *Entscheidungsproblem*", *Journal of Symbolic Logic,* I (1936), 40-41, 101-102. Esta nota es breve, pero depende de una buena comprensión de los resultados relativos al denominado por Church "Calculus of Lambda-Conversion". Este es equivalente a la teoría de la computabilidad (máquinas Turing); y una prueba del teorema de Church desde este último punto de vista (que resultará fácilmente comprensible para la mayoría de los lectores) se encuentra en la obra de Martin Davis (ed.), *Computability and Unsolvability* (New York: McGraw-Hill, 1958), capítulo 8, sección 8. Esta sección no puede ser comprendida por sí misma, ya que depende de los resultados de la parte anterior del libro.
3. A. Tarski, *A Decision Method for Elementary Algebra and Geometry,* 2.ª edición (Berkeley and Los Angeles: RAND Corporation, 1951).
4. Kurt Gödel, *The Consistency of the Axiom of Choice and the Generalized*

Continuum-Hypothesis with the Axioms of Set Theory (Princeton: Princeton University Press, 1940).
5. Paul J. Cohen, "The Independence of the Continuum Hypothesis, "*Proceedings of the National Academy of Sciences* (U. S. A.), L (1963), 1143-1148, y LI (1964), 105-110.
6. A. Tarski, "*Der Wahrheitsbegriff in den formalisierten Sprachen*", *Studia Philosophica,* I (1936), 261-405. Traducción inglesa de A. Tarski, en *Logic, Semantics, Metamathematics* (Oxford: Oxford University Press, 1956).
7. Véase W. V. Quine, *Mathematical Logic,* edición revisada (New York: Holt, Rinehart & Winston, 1959), sección 5.
8. K. R. Popper, "The Propensity Interpretation of the Calculus of Probability and the Quantum Theory", en S. Körner (ed.), *Observation and Interpretation in the Philosophy of Physics* (New York: Dover, 1962), y "The Propensity Interpretation of Probability", *British Journal for the Philosophy of Science,* X (1959), 25-42.

CAPÍTULO 3:

1. Véase el informe en *Scientific American* (julio, 1965), pág. 46.
2. C. G. Hempel, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (New York: Free Press, 1965).
3. Michael Scriven, "Explanations, Predictions and Laws", en H. Feigl y G. Maxwell (eds.), *Scientific Explanation, Space and Time, Minnesota Studies in the Philosophy of Science,* Vol. III (Minneapolis: University of Minnesota Press, 1962), págs. 170-230.
4. Hempel, *op. cit.*
5. Nelson Goodman, *Fact, Fiction and Forecast* (Cambridge: Harvard University Press, 1955), capítulo 3.
6. A. J. Ayer, "Review of Ernest Nagel's *Structure of Science*", *Scientific American* (junio, 1961), 197-203.
7. Véase Alan Moorehead, *The White Nile* (London: Hamish Hamilton, 1960), página 47.
8. W. Dray, *Laws and Explanation in History* (Oxford: Oxford University Press, 1957).
9. Hempel, *op. cit.,* págs. 415-425.
10. *Ibíd.,* pág. 421, nota 14.
11. Véase especialmente N. R. Campbell, *Physics, The Elements* (London: Cambridge University Press, 1920), reeditada como *Foundations of Science* (New York: Dover, 1957).
12. Campbell, *Foundations of Science,* pág. 123.
13. *Ibíd.,* pág. 144.
14. Hempel, *op. cit.,* págs. 430-431.
15. G. Schlesinger, "What Is Science For?", *Australian Journal of Science,* XXVI (1963), 163-167.
16. Véase especialmente P. K. Feyerabend, "Explanation, Reduction and Empiricism", en H. Feigl y G. Maxwell (eds.), *Scientific Explanation, Space and Time, Minnesota Studies in the Philosophy of Science,* Vol. III (Minneapolis: University of Minnesota Press, 1962)), págs. 28-97; "Reply to Criticism, Comments on Smart, Sellars and Putnam", en R. S. Cohen y M. W. Wartofsky (eds.), *Boston Studies in the Philosophy of Science,* Vol. II (New York: Humanities Press, 1965), págs. 223-261; y "Problems of Empiricism", en R. G. Colodny (ed.), *Beyond the Edge of Certainty* (Pittsburgh: University of Pittsburgh Press, 1965).
17. J. Kemeny y P. Oppenheim, "On Reduction", *Philosophical Studies,* VII (1956), 6-19.
18. Hempel, *op. cit.,* pág. 347, nota a pie de página.

19. Ernest Nagel, *The Structure of Science* (New York: Harcourt, Brace and World, 1961), capítulo 11.
20. Hilary Putnam, "How Not to Talk about Meaning, Comments on J. J. C. Smart", en R. S. Cohen y M. W. Wartofsky (eds.), *Boston Studies in the Philosophy of Science,* Vol. II (New York: Humanities Press, 1965), página 207.
21. Véase Wilfrid Sellars, *Science, Perception and Reality* (New York: Humanities Press, 1963), capítulo 4, traducción española: *Ciencia, percepción y realidad,* ed. Tecnos, Madrid, 1971; y "Scientific Realism or Irenic Instrumentalism, Comments on J. J. C. Smart", en R. S. Cohen y M. W. Wartofsky (eds), *Boston Studies in the Philosophy of Science,* Vol. II (New York: Humanities Press, 1965), págs. 171-204.
22. Véase, por ejemplo, Sir Arthur Eddington, *The Philosophy of Physical Science* (London: Cambridge University Press, 1939).
23. Feyerabend, "Explanation, Reduction and Empiricism", *op. cit.,* pág. 260, nota 44.
24. R. Carnap, "Testability and Meaning", *Philosophy of Science,* III (1936), 419-471; y IV (1937), 1-40.
25. Feyerabend, "Explanation, Reduction and Empiricism", *op. cit.,* págs. 65-66.
26. Véase A. Einstein, *Investigations on the Theory of the Brownian Movement,* traducido por A. D. Cooper, y ed. con notas por R. Fürth (New York: Dover, 1956).
27. Feyerabend, "Explanation, Reduction and Empiricism", *op. cit.,* pág. 37.
28. Putnam, *op. cit.,* pág. 210.
29. Véase David Lewis, "An Argument for the Identity Theory", *Journal of Philosophy,* LXIII (1966), 17-25.
30. E. B. Uvarov, D. R. Chapman y Alan Isaacs, *A Dictionary of Science,* 3.ª ed. (Baltimore, Md.: Penguin Books, 1964).
31. Feyerabend, "Reply to Criticism", *op. cit.,* pág. 248; y "Problems of Empiricism", *op. cit.,* pág. 256.
32. Véase Putnam, *op. cit.,* pág. 211.
33. *Ibíd.,* pág. 216.
34. Sir George Thomson, "Some Thoughts on the Scientific Method", en R. S. Cohen y M. W. Wartofsky (eds.), *Boston Studies in the Philosophy of Science,* Vol. II (New York: Humanities Press, 1965).

CAPÍTULO 4:

1. Robert R. Sokal, "The Future Systematics" en Charles A. Leone (ed.), *Taxonomic Biochemistry and Serology* (New York: Ronald Press, 1964); y "Numerical Taxonomy", *Scientific American* (diciembre, 1966), 106-116.
2. Véase Michael Scriven, "Truisms as the Grounds for Historical Explanations" en P. Gardiner (ed.), *Theories of History* (New York: Free Press, 1959), páginas 443-475.
3. Véase la respuesta de C. G. Hempel a Scriven en relación con el caso del puente que se hunde a causa del envejecimiento del metal en su *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (New York: Free Press, 1965), pág. 371.
4. Véase R. A. Fisher, *The Genetical Theory of Natural Selection,* 2.ª ed. (New York: Dover, 1958).
5. Véase G. G. Simpson, *The Meaning of Evolution* (New York: Mentor Books, 1954), págs. 34-35.
6. Michael Scriven, "Explanation and Prediction in Evolutionary Theory", *Science,* CXXX (1959), 477-482.
7. Sobre las máquinas auto-reproductoras, véase J. von Neumann, "The General and Logical Theory of Automata", en L. A. Jefress (ed.), *Cerebral Mecha-*

nisms in Behavior, The Hixon Symposium (New York: Wiley, 1951), páginas 1-41.
8. Véase W. T. Williams, "Problems of Alien Biology", *Humanist,* LXXIX (noviembre, 1964), 329-332.
9. *Ibíd.*
10. L. J. Henderson, *The Fitness of the Environment* (New York: Macmillan, 1913).
11. J. A. Deutsch, *The Structural Basis of Behavior* (London: Cambridge University Press, 1960).
12. Jerry A. Fodor, "Explanations in Psychology" en Max Black (ed.), *Philosophy in America* (London: Allen & Unwin, 1965, e Ithaca: Cornell University Press, 1965), págs. 161-179.
13. B. A. Farrell, "On the Limits of Experimental Psychology", *British Journal of Psychology,* LVI (1955), 165-177.
14. Deutsch, *op. cit.,* pág. 12 y capítulo 10.
15. *Ibíd.,* págs. 14-15.
16. Fodor, *op. cit.*
17. *Ibíd.,* pág. 179.

CAPÍTULO 5:

1. N. R. Campbell, *An Account of the Principles of Measurement and Calculation* (London: Longmans, Green, 1928); y *Foundations of Science* (New York: Dover, 1957).
2. Véase Campbell, *Foundations of Science, op. cit.,* pág. 271.
3. Sir Arthur Eddington, *The Nature of the Physical World* (London: Dent, 1964), pág. 245.
4. Véase N. R. Campbell, "The Errors of Sir Arthur Eddington", *Philosophy,* VI (1930), 180-192.
5. Michael Scriven, "The Key Property of Physical Laws-Inaccuracy", en H. Feigl y G. Maxwell (eds.), *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart & Winston, 1961), págs. 91-101; véase, especialmente, la pág. 101.
6. *Ibíd.*
7. Véanse los comentarios de H. Mehlberg sobre el trabajo de Scriven, también en H. Feigl y G. Maxwell (eds.), *Current Issues in the Philosophy of Science* New York: Holt, Rinehart & Winston, 1961), págs. 102-103.
8. M. Scriven, "Rejoinder to Mehlberg", en H. Feigl y G. Maxwell (eds.), *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart & Winston, 1961), págs. 103-104.
9. *Ibíd.,* pág. 104.
10. Véase, especialmente, la obra del obispo George Berkeley, *Principles of Human Knowledge,* especialmente las secciones 58-66 y las secciones 101-117; y su *De Motu.* Véase David M. Armstrong (ed.), *Berkeley's Philosophical Writings* (New York: Collier, 1965).
11. Ernest Mach, *Contributions towards the Analysis of Sensations* (Chicago: Open Court, 1897); y *The Science of Mechanics: a Critical and Historical Account of its Development,* trad. por Thomas J. McCormack, 6.ª ed. revisada (La Salle, Ill.: Open Court, 1960), especialmente págs. 577-595.
12. Wilfrid Sellars, *Science, Perception and Reality* (London: Routledge & Kegan Paul, 1963), capítulo 3 (traducción española: *Ciencia, percepción y realidad,* ed. Tecnos, Madrid, 1971).
13. Véase P. W. Bridgman, "Operational Analysis", *Philosophy of Science,* V (1938), 114-131, especialmente la pág. 114.
14. P. W. Bridgman, "Einstein's Theories and the Operational Point of View",

en P. A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist,* 2.ª ed. (La Salle, Ill.: Open Court, 1951), págs. 333-354.

15. Albert Einstein, "On the Electrodynamics of Moving Bodies", en *The Principle of Relativity, a Collection of Original Memoirs on the Special and General Theory of Relativity,* por H. A. Lorentz, A. Einstein, H. Minkowski y H. Weyl, con notas a cargo de A. Sommerfeld, trad. por W. Perrett y G. B. Jeffrey (London: Methuen, 1923).
16. Véase, por ejemplo, W. Rindler, *Special Relativity* (Edinburgh: Oliver and Boyd, 1960), pág. 12.
17. Albert Einstein, "Autobiographical Notes", en P. A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist,* 2.ª ed. (La Salle, Ill.: Open Court, 1951). La influencia de Hume y Mach en él puede verse en la pág. 53; su posterior rechazo de la posición de Mach puede verse en la pág. 21.
18. Como hace N. R. Hanson. Véase su "Number Theory and Physical Theory: An Analogy", en R. S. Cohen y M. W. Wartofsky (eds.), *Boston Studies in the Philosophy of Science,* Vol. II (New York: Humanities Press, 1965).
19. Gibert Ryle, *The Concept of Mind* (New York: Barnes & Noble, 1949), páginas 120-125.
20. Stepehn Toulmin, *The Philosophy of Science* (London: Hutchinson, 1953).
21. Véase F. P. Ramsey, *Foundations of Mathematics* (London: Kegan Paul, 1931), página 231.
22. Véase W. V. Quine, *Set Theory and its Logic* (Cambridge: Harvard University Press, 1963), pág. 231.
23. William Craig, "On Axiomatizability within a System", *Journal of Symbolic Logic,* XVIII (1953), 30-32; y "Replacement of Auxiliary Expressions", *Philosophical Review,* LXV (1956), 38-55.
24. Véase Marshall Spector, "Models and Theories", *British Journal for the Philosophy of Science,* XVI (1965), 121-142.
25. C. G. Hempel, "The Theoretician's Dilemma", en H. Feighl, M. Scriven y G. Maxwell (eds.), *Scientific Explanation, Space and Time, Minnesota Studies in the Philosophy of Science,* Vol. II (Minneápolis: University of Minnesota Press, 1958), págs. 37-98, véase pág. 79.
26. Grover Maxwell, "The Ontological Status of Theoretical Entities", en H. Feigl y G. Maxwell (eds.), *Scientific Explanation, Space and Time, Minnesota Studies in the Philosophy of Science,* Vol. III (Minneapolis: University of Minnesota Press, 1962), págs. 3-17, véase la pág. 9.
27. Véase P. K. Feyerabend, "Explanation, Reduction and Empiricism", en H. Feigl y G. Maxwell (eds.), *Scientific Explanation, Space and Time, Minnesota Studies in the Philosophy of Science,* Vol. III (Minneapolis: University of Minnesota Press, 1962), págs. 28-97. Véanse las págs. 36-39.
28. R. Carnap, "Testability and Meaning", *Philosophy of Science,* III (1936), 419-471; y IV (1937), 1-40.
29. K. R. Popper, *Conjectures and Refutations* (London: Routledge & Kegan Paul, 1963), capítulo 3, especialmente las págs. 111-114. El instrumentalismo también ha sido criticado en los demás escritos de Popper.
30. P. K. Feyerabend, "Realism and Instrumentalism" en Mario Bunge (ed.), *The Critical Approach to Science and Philosophy* (New York: Free Press, 1964), páginas 280-308.
31. Véase Bernard Mayo, "More about Theoretical Entities", *Penguin Science News,* núm. 39 (febrero, 1956), 42-55, especialmente el parágrafo final.
32. K. R. Popper, "The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory", en S. Körner (ed.), *Observation and Interpretation in the Philosophy of Physics* (New York: Dover, 1962), págs. 65-70, y la discusión de las págs. 78-89.
33. P. K. Feyerabend, "Problems of Microphysics" en R. G. Colodny (ed.), *Fron-*

tiers of Science and Philosophy (Pittsburgh: University of Pittsburgh Press, 1962), véase pág. 200.

34. H. Margenau, *The Nature of Physical Reality* (New York: McGraw-Hill, 1950), página 435 (traducción española: *La naturaleza de la realidad física,* ed. Tecnos, Madrid, 1970).
35. H. Margenau, *Open Vistas* (New Haven, Conn.: Yale University Press, 1961), véase la pág. 161.
36. J. C. C. McKinsey, A. C. Sugar y P. Suppes, "Axiomatic Foundations of Particle Mechanics", *Journal of Rational Mechanics and Analysis,* II (1953), páginas 253-277.
37. W. V. Quine, *Word and Object* (Cambridge: M. I. T. Press, 1959), capítulo 6, especialmente la sección 4-1; y "Three Grades of Modal Involvement", *Proceedings of XIth International Congress of Philosophy,* XIV (Bruselas, 1953), 65-81, reeditado en W. V. Quine *The Ways of Paradox* (New York: Random House, 1966), págs. 156-174.
38. Véase la explicación de Rudolf Carnap de su lógica modal en P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (La Salle, Ill.: Open Court, 1963), páginas 889-900; también sus observaciones en la pág. 63 del mismo volumen; y su *Meaning and Necessity* (Chicago: University of Chicago Press, 1949). Véase, por ejemplo, el simposio entre Ruth Barcan Marcus y W. V. Quine, junto con la discusión posterior, en M. W. Wartofsky (ed.), *Boston Studies in the Philosophy of Science,* Vol. I (Dordrecht, Holland: D. Reidel, 1963), páginas 77-116. La réplica de Quine a Marcus ha sido reeditada en su *The Ways of Paradox* (New York: Random House, 1966), págs. 175-182.

CAPÍTULO 6:

1. Véase David Hume, *Treatise on Human Nature,* libro I, parte 3, sección 6, y *Enquiry Concerning Human Understanding,* sección IV, parte 2.
2. Max Black, *Problems of Analysis* (Ithaca: Cornell University Press, 1954), capítulo 11; y *Models and Metaphors* (Ithaca: Cornell University Press, 1962), capítulo 12 (traducción española: *Modelos y metáforas,* ed. Tecnos, Madrid, 1967).
3. R. B. Braithwaite, *Scientific Explanation* (London: Cambridge University Press, 1953), capítulo 8 (traducción española: *La explicación científica,* ed. Tecnos, Madrid, 339).
4. J. M. Keynes, *A. Treatise on Probability* (London: Macmillan, 1921), capítulo 22.
5. Paul Edwards, "Bertrand Russell's Doubts about Induction" en A. G. N. Flew (ed.), *Logic and Language, First Series* (Oxford: Blackwell, 1951), págs. 55-79. Edwards comenta el capítulo 6 de la obra de Russell, *Problems of Philosophy* (Oxford: Oxford University Press, 1946).
6. Edwards, *op. cit.,* pág. 59.
7. Herbert Feigl, "*De Principiis Non Disputandum...*?" en Max Black (ed.), *Philosophical Analysis* (Ithaca: Cornell University Press, 1950), págs. 119-156.
8. Hans Reichenbach, "On the Justification of Induction", *Journal of Philosophy,* XXXVII (1940), págs. 97-103. Véase también la sección 39 de su *Experience and Prediction* (Chicago: University of Chicago Press, 1938).
9. Antony Flew, "Is Pascal's Wager the Only Safe Bet?", *Rationalist Annual* (1960), 21-25.
10. James Cargile, "Pascal's Wager", *Philosophy,* XLI (1966), 250-257.
11. Isabel Creed, "The Justification of the Habit of Induction", *Journal of Philosophy,* XXXVII (1940), 85-97.
12. Jerrold J. Katz, *The Problem of Induction and its Solution* (Chicago: University of Chicago Press, 1962).
13. Véase, en particular, Wesley C. Salmon, "Inductive Inference", en Bernard

Baumrin (ed.), *Philosophy of Science, The Delaware Seminar,* Vol. 2 (New York: Interscience Publishers, 1963), págs. 341-370. así como su trabajo "Vindication of Induction", con comentarios de Stephen Barker, con réplica de Salmon, y comentarios de Richard Rudner, en H. Feigl y G. Maxwell (eds.), *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart & Winston, 1959), págs. 245-264.

14. Stephen Barker, "Comment on 'Vindication of Induction'", en H. Feigl y G. Maxwell (eds.), *Current Issues in the Philosophy of Science* (New York: Holt, Rinehart & Winston, 1959).
15. Nelson Goodman, *Fact, Fiction and Forecast* (Cambridge: Harvard University Press, 1955), capítulo 3.
16. Wesley C. Salmon, "The Foundations of Scientific Inference", en R. C. Colodny (ed.), *Mind and Cosmos* (Pittsburgh: University of Pittsburgh Press, 1966), págs. 135-275. Sobre el rechazo de los predicados de observación básicos véase la pág. 157; una referencia a Ian Hacking puede verse en la página 239.
17. Ian Hacking, "Salmon's Vindication of Induction", *Journal of Philosophy,* LXII (1965), 260-266.
18. Véase Katz, *op. cit.,* pág. 61.
19. Wesley C. Salmon, "The Short Run", *Philosophy of Science,* XXII (1955), 214-221.
20. Brian Ellis, "A Vindication of Scientific Inductive Practices", *American Philosophical Quarterly,* II (1965), 296-304.
21. K. R. Popper, *Conjectures and Refutations* (London: Routledge & Kegan Paul, 1963), pág. 47.
22. Hilary Putnam, "'Degree of Confirmation' and Inductive Logic", en P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (La Salle, Ill.: Open Court, 1963), págs. 761-783, véanse, especialmente, las págs. 770-777.
23. R. B. Braithwaite, *op. cit.*
24. Pierre Duhem, *The Aim and Structure of Physical Theory* (Princeton: Princeton University Press, 1954).
25. Ellis, *op. cit.,* págs. 304.
26. Véase, por ejemplo, la obra de Popper, *Conjectures and Refutations, op. cit.,* capítulo 1, especialmente págs. 53-55.
27. Véase Rudolf Carnap, *Logical Foundations of Probability,* 2.ª ed. (Chicago: University of Chicago Press, 1962) y *The Continuum of Inductive Methods* (Chicago: University of Chicago Press, 1952). Véase también la explicación de Carnap de sus conceptos básicos de probabilidad e inducción en las páginas 966-979 de su "Replies and Expositions", en P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (La Salle, Ill.: Open Court, 1963).
28. Popper, *Conjectures and Refutations, op. cit.,* pág. 35.
29. K. R. Popper, *The Logic of Scientific Discovery* (London: Routledge & Kegan Paul, 1959), véanse las págs. 375 y siguientes (traducción española: *La lógica de la investigación científica,* ed. Tecnos, Madrid, 1962).
30. Popper, *Conjectures and Refutations, op. cit.,* pág. 242.
31. Adolf Grünbaum, *Philosophical Problems of Space and Time* (New York: Knopf, 1963).

CAPÍTULO 7:

1. Bertrand Russell, *The Philosophy of Leibniz* (London: Allen & Unwin, 1900).
2. Las citas del texto son de la edición de Florian Cajori de *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World* (Berkeley: University of California Press, 1934).
3. Ernest Mach, *The Science of Mechanics: A Critical and Historical Account*

of its Development, trad. de Thomas J. McCormack, 6.ª ed. rev. (La Salle, Ill.: Open Court, 1960), especialmente el capítulo 2, sección VI, números 2-6.
4. Gilbert Ryle, "The Verification Principle", *Revue Internationale de Philosophie,* III-IV (1951), 243-250.
5. Innanuel Kant, *Prolegomena to Any Future Metaphysics,* trad. con introducción y notas, de P. G. Lucas (Manchester: Manchester University Press, 1953).
6. Adolf Grünbaum, *Philosophical Problems of Space and Time* (New York: Knopf, 1963), págs. 386-397.
7. H. Minkowski, "Space and Time", en *The Principle of Relativity, A Collection of Original Memoirs on the Special and General Theory of Relativity,* por H. A. Lorentz, A. Einstein, H. Minkowski y H. Weyl, con notas de A. Sommerfeld, trad. por W. Perrett y G. B. Jeffrey (London: Methuen, 1923), páginas 73-91, véanse especialmente págs. 78-79.
8. *Ibíd.,* pág. 75.
9. Véase Moritz Schlick, *Philosophy of Nature,* trad. por Amethe von Zeppelin (New York: Philosophical Library, 1949), pág. 43.
10. Véase E. A. Milne y G. J. Whitrow, "On the so-called 'Clock-paradox' of Special Relativity", *Philosophical Magazine,* XL (1949), 1244-1249; y G. J. Whitrow, *The Natural Philosophy of Time* (London: Nelson, 1961), páginas 219-222.
11. Véase Grünbaum, *op. cit.,* capítulo 14.
12. *Ibíd.,* pág. 419.
13. J. A. Wheeler, "Curved Empty Space-Time as the Building Material of the Physical World: an Assessment", en E. Nagel, P. Suppes y A. Tarski (eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the* 1960 *International Congress* (Standford: Standford University Press, 1962), páginas 361-374.
14. Hans Reichenbach, *The Philosophy of Space and Time,* trad. de Maria Reichenbach y John Freud, con introducción a cargo de Rudolf Carnap (New York: Dover, 1958), sección 12.
15. Véase Grünbaum, *op. cit.,* pág. 97.
16. *Ibíd.,* capítulo 4.
17. Hilary Putnam, "An Examination of Grünbaum's Philosophy of Geometry", en B. Baumrin (ed.), *Philosophy of Science, The Delaware Seminar,* Vol. 2 (New York: Interscience, 1963), págs. 205-255.
18. *Ibíd.,* págs. 219-220.
19. Ernest Nagel, *The Structure of Science* (New York: Harcourt, Brace & World, 1961), capítulo 9.
20. Adolf Grünbaum, *op cit.,* capítulo 6. Grünbaum ha dado un tratamiento más extenso en su trabajo "A Consistent Conception of the Extended Linear Continuum as an Aggregate of Unextended Elements", *Philosophy of Science,* XIX (1952), 288-306, y en el capítulo 3 de su libro *Modern Science and Zeno's Paradoxes* (Middletown, Conn.: Wesleyan University Press, 1967).
21. W. V. Quine, *Word and Object* (Cambridge: M. I. T. Press, 1960).

CAPÍTULO 8:

1. Moritz Schlick, *Philosophy of Nature,* trad. de Amethe von Zeppelin (New York: Philosophical Library, 1949).
2. Adolf Grünbaum, *Philosophical Problems of Space and Time* (New York: Knopf, 1963).
3. Grünbaum, "Carnap's Views on the Foundations of Geometry", en P. A. Schilpp (ed.), *The Philosophy of Rudolf Carnap* (La Salle, Ill.: Open Court, 1963).
4. J. H. Christenson, J. W. Cronin, V. L. Fitch y R. Turlay, "Evidence for

the 2 π Decay of the K_2^0 Meson", *Physical Review Letters,* XIII (1964), 138-140.
5. J. Loschmidt, "Uber das Wärmegleichgewicht eines Systems von Körpen mit Rücksicht auf die Schwere", *Sitzungsberichte der Akademie der Wissenschaften in Wien,* LXXIII (1876), y LXXXV (1877).
6. S. Toulmin, R. W. Hepburn, y A. MacIntyre, *Methaphysical Beliefs* (London: S. C. M. Press, 1957).
7. F. H. Bradley, *Appearance and Reality,* 2.ª ed. (Oxford: Oxford University Press, 1930), pág. 190.
8. Claude E. Shannon, "The Mathematical Theory of Communication", originalmente publicado en el *Bell System Technical Journal,* XXVII (1948), 379-423 y 623-656; y reeditado en Claude E. Shannon y Warren Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1949).
9. Y. Bar-Hillel y R. Carnap, "Semantic Information", *British Journal for the Philosophy of Science,* XIV (1953-1954), 147-157.
10. Hans Reichenbach, *The Direction of Time* (Berkeley: University of California Press, 1956), págs. 156 y siguientes.
11. A. Grünbaum, *Philosophical Problems of Space and Time, op. cit.,* pág. 258.
12. Reichenbach, *op. cit.*
13. Max Black, *Models and Metaphor* (Ithaca: Cornell University Press, 1962), capítulo 10, "The 'Direction' of Time". Véase especialmente la parte final, a partir de la mitad, de la pág. 191 del ensayo de Black (traducción española: *Modelos y metáforas,* ed. Tecnos, Madrid 344).
14. Grünbaum, *Philosophical Problems of Space and Time, op. cit.,* pág. 195.
15. Véase, especialmente, Grünbaum, *ibíd.,* pág. 258, parágrafo final.
16. *Ibíd.,* pág. 257.
17. Norbert Wiener, *Cybernetics,* 2.ª ed. (New York: Wiley, 1961).
18. Véase, por ejemplo, Charles A. Coulson, *Electricity* (New York: Interscience, 1948), sección 113.
19. O. Costa de Beauregard, *Le Second Principe de la Sciende du Temps* (París: Editions du Seuil, 1963), págs. 33-37.
20. K. R. Popper, *Letters in Nature,* CLXXVII (1956), 538; CLXVIII (1956), 382; LXXIX (1957), 1297; LXXXI (1958), 402-403. Críticas a Popper pueden verse en A. Grünbaum, "Popper on Irreversibility" en Matio Bunde (ed.), *The Critical Approach to Science and Philosophy, in Honor of Karl Popper* (New York: Free Press, 1964, págs. 316-331.
21. Sir A. S. Eddington, *The Expanding Universe* (London: Cambridge University Press, 1933), pág. 66.
22. T. Gold, "The Arrow of Time" en *La Structure et L'Evolution de l'Univers, Proceedings of the 11th Solvay Conference* (Bruselas: R. Stoops, 1958), 81-91; y otro trabajo con el mismo título en *American Journal of Physics,* XXX (1962), 403-410.
23. Véase H. Bondi, *Cosmology,* 2.ª ed. (London: Cambridge University Press, 1960), pág. 22.
24. Gold, "The Arrow of Time", *op cit.* (1958), pág. 85.
25. *Ibíd.,* pág. 82.
26. T. Gold, "Cosmic Processes and the Nature of Time", en R. G. Colodny (ed.), *Mind and Cosmos* (Pittsburgh: University of Pittsburgh Press, 1966), páginas 311-329.
27. H. Bondi, "Physics and Cosmology", *The Observatory,* LXXXII (1962), 133-143.
28. *Ibíd.,* págs. 142-143.

CAPÍTULO 9:

1. Marquis P. S. de Laplace, *A Philosophical Essay on Probabilities,* trad. de la 6.ª ed. francesa por F. W. Truscott y F. L. Emory (New York: Dover, 1951), pág. 4.

2. K. R. Popper, "Indeterminism in Quantum Physics and in Classical Physics", *British Journal for the Philosophy of Science*, I (1950-1951), 117-133 y 173-195.
3. G. Ryle, *The Concept of Mind* (New York: Barnes & Noble, 1949), págs. 76-81.
4. G. F. Dear, "Determinism in Classical Physics", *British Journal for the Philosophy of Science*, XI (1960-61), 289-304.
5. *Ibíd.*, pág. 293.
6. D. W. Sciama, "Determinism and the Cosmos", en Sidney Hook (ed.), *Determinism and Freedom in the Age of Modern Science* (New York: Collier, 1961), páginas 90-91.
7. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, trad. por R. T. Beyer (Princeton: Princeton University Press, 1955), capítulo 4, sección 2.
8. Sciama, *op. cit.*, pág. 91.
9. J. C. Eccles, "Hypotheses Relating to the Brain-Mind Problem", *Nature*, CLXVIII (1951), 53-56.
10. D. M. MacKay, "Brain and Will", *Listener*, LVII (1957), 788-789.
11. W. Ross Ashby, "Design for an Intelligence Amplifier", en C. E. Shannon y J. McCarthy (eds.), *Automata Studies, Annals of Mathematics Studies*, núm. 34 (Princeton University Press, 1956), págs. 215-234.
12. Alfred North Whitehead y Bertrand Russell, *Principia Mathematica*, 1.ª ed. (Cambridge University Press, 1913).
13. Véase D. Hilbert y W. Ackermann, *Principles od Mathematical Logic*, trad. de L. M. Hammond, G. G. Leckie y F. Steinhardt, y ed. con notas por R. E. Luce (New York: Chelsea, 1950, pág. 27.
14. Alfred North Whitehead y Bertrand Russell, *Principia Mathemathica*, 2.ª ed. (London: Cambridge University Press, 1927), *1.2 a *1.6.
15. P. Bernays, "*Axiomatische Untersuchung des Aussagenkalküls der Principia Mathematica*", *Mathematische Zeitschrift*, XXV (1926), 305-320.
16. A. Newell, J. C. Shaw y H. A. Simon, "Theory of Human Problem Solving", *Psychological Review*, LXV (1958), 151-166.
17. Newell, Shaw y Simon, *op. cit.*, pág. 158.
18. Noam Chomsky, "Review of B. F. Skinner's *Verbal Behavior*", *Language*, XXXV (1959), 26-58. Esta revisión ha sido reeditada en Jerry A. Fodor y Jerrold J. Katz (eds.), *The Structure of Language, Readings in the Philosophy of Language* (Englewood Cliffs, N. J.: Prentice-Hall, 1964), págs. 547-578.
19. B. F. Skinner, *Verbal Behavior* (New York: Appleton-Century-Crofts, 1957).
20. H. L. Gelernter "Realization of a Geometry Proving Machine", *Proceedings of the International Conference on Information Processing* (París: UNESCO House, 1959), págs. 273-282.
21. Marvin L. Minsky, "Some Methods of Artificial Intelligence and Heuristic Programming", National Physical Laboratory Symposium on *Mechanisation of Thought Processes* (London: Her Majesty's Stationery Office, 1959), Vol. I, páginas 5-27.
22. Véase, por ejemplo, J. R. Lucas, "Minds, Machines and Gödel", *Philosophy*, XXXVI (1961), 112-127. Este ha sido reeditado en A. R. Anderson (ed.), *Minds and Machines* (Englewood Cliffs, N. J.: Prentice-Hall, 1964), 43-59.
23. Hilary Putnam, "Minds and Machines", en Sidney Hook (ed.), *Dimensions of Mind* (New York: Collier, 1961), 138-164.
24. G. Gentzen, "Die Widerspruchsfreiheit der reinen Zahlentheorie", *Mathematische Annalen*, CXII (1936), 493-565.
25. W. Grey Walter, *The Living Brain* (London: Duckworth, 1953).
26. Richard Taylor, *Action and Purpose* (Englewood Cliffs, N. J.: Prentice-Hall, 1966), capítulo 15.
27. Véase la discusión del ejemplo de Taylor realizada por D. M. Armstrong en su nota crítica a la obra de Taylor, *Action and Purpose, Australasian Journal of Philosophy*, XLIV (1966), 231-240, especialmente págs. 238-239.

28. G. Sommerhoff, *Analytical Biology* (Oxford: Oxford University Press, 1950), páginas 85 y siguientes.
29. W. Ross Ashby, *An Introduction to Cybernetics* (New York: Wiley, 1961).
30. Y. Bar-Hillel, *Language and Information* (Reading, Mass.: Addison-Wesley, 1964), pág. 179.
31. W. E. Bull, C. Africa y D. Teichrow, "Some problems of the 'word'", en W. N. Locke y A. D. Booth (eds.), *Machine Translation* (Cambridge: Massachusetts Institute of Technology Press, 1955), págs. 86-103.
32. N. Chomsky, *Syntactic Structures* (The Hague: Mouton, 1957).
33. C. P. Snow, *The Two Cultures and the Scientific Revolution* (London: Cambridge University Press, 1959).
34. P. F. Strawson, *Individuals* (London: Methuen, 1959).

INDICE ALFABETICO

A

Achinstein, 72.
Ackermann, 45, 234 (nota).
Adán, 88.
Adams, J. C., 153.
Africa, C., 245.
Agassi, Joseph, 133.
Alexander, H. Gavin, 133, 193.
Algoritmo, 30, 115-116.
Ambito (Scope), 28.
Analogía, 59-62.
Análisis, 34-35, 191.
Anderson, A. R., 248.
Aristóteles, 125, 162.
Ashby, W. Ross, 232, 245 (nota), 247.
Astronomía, 15, 74, 128-129, 149, 151, 196.
Ayer, A. J., 73.

B

Bar-Hille, Y., 203 (nota), 218-219, 245.
Barker, Stephen, 144, 155.
Baumrin, B., 72-73, 155, 193.
Beauregard, O. Costa de, 192, 210, 218.
Bellarmine, Cardenal, 121.
Benacerraf, P., 248.
Bennett, Jonathan, 155.
Bergmann, Gustav, 134.
Berkeley, George, 108, 193.
Bernays, P., 234 (nota).
Billete de inferencia, 133.
Black, Max, 95, 137, 156, 207, 218.
Bohr, Niels, 61-62.
Boltzmann, L., 201, 218.
Bolyai, W., 189.
Bondi, H., 217 (nota), 219.
Borel, E., 192.
Born, Max, 133.
Bradley, F. H., 200 (nota).
Braithwaite, R. B., 24, 95, 137, 248.
Brasil, 151.
Bridgman, P. W., 109, 133.
Broad, C. D., 155.
Brodbeck, M., 24, 73.
Bromberger, Sylvian, 73.
Brujería, 22.
Bull, W. E., 245.
Bunge, Mario, 95.
Burks, A. W., 156.
Butler, R. J., 156.
Butts, Robert E., 73.

C

Campbell, C. A., 247.
Campbell, N. R., 59, 100-103, 132.
Cargile, James, 141 (nota).
Carnap, Rudolf, 24, 46, 67 (nota), 113 (nota), 120, 128, 150-151, 156, 203, (nota), 219.
Casualidad, 224-225, 226.
Cherry, Colin, 218.
Chomsky, Noam, 17, 95, 237, 246, 248.
Christenson, J. H., 197.
Chruch, Alonzo, 30-31, 45.
Churchman, C. W., 132.
Cibernética, 14, 241-246.
Clarke, Samuel, 193.
Cohen, E. G. D., 218.
Cohen, Paúl J., 34.
Cohen, R. S., 72, 248.
Colodny, R. G., 25, 72-73, 155, 157, 219, 247.
Completitud, 29.
Computabilidad, 45.
Conducta, 90-91.
Condicional fuerte, 127, 129-131, 221.
Confirmación, 150-151.
Confirmación de muestra, 150.
Contador de Geiger, 225, 229.
Contexto extensional, 37-38, 127.
Contexto intensional, 37, 38.
Convención de congruencia, 186 y ss.
Copérnico, 118, 121.
Corroboración, 151-152.

Cosmología, 74, 131, 164-165, 179-183, 200, 206, 212-214, 215-219.
Craig, William, 115, 116-118, 133.
Creed, Isabel, 141.
Creencia racional, 41.
Cristianismo, 141.
Crommelin, A. C. D., 151.
Cronin, J. W., 197.
Cuantificadores, 28.
Cubo de Newton, 161-162.
Curvatura, 181-183, 190.

D

Danto, A., 25, 133.
Darwin, Charles, 81-82.
Davidson, C., 151.
Davis, Martin, 45.
Dear, G. F., 222.
Demócrito, 22.
Denotación, 29.
Dependencia de contexto, 36, 194.
Descartes, René, 158, 164, 182-183.
Descripción, 31-32, 51.
Determinismo, 173, 220-230.
Deutsch, J. A., 90 (nota), 92-95.
"Dib" (Digitos binarios), 202-203.
Dios, 81-82, 164.
Dray, William, 57.
Duhem, Pierre, 149, 153-154, 157.

E

Eccles, J. C., 229.
Ecología, 85-86.
Economía, 16.
Eddington, A. S., 66, 103, 151-152, 192, 211.
Edwards, Paúl, 139-141, 155, 156.
Eistein, Albert, 23, 70, 105, 109-110, 151-152, 162, 168, 182, 192.
Electrónica, 79-81, 89-91, 94.
Ellis, Brian, 132, 146-148, 149, 155.
Entropía, 64, 201-210.
Enumeración, simple, 145, 147.
Escala de Mohs, 100.
Espacio, 158-165.
Espacio absoluto, 158-165, 179.
Espacio relativo, 158-159, 162-163, 164-165, 192.
Espacio-tiempo, 183-189, 192-195.
Espacio-tiempo absoluto, 164, 179-183.
Espacio-tiempo relativo, 180-183.
Esquema, 27-29, 191.
Esbozo de explicación, 57, 59, 83.
Esquema tautológico, 27-28.
Esquema válido, 27-29.
Estrategias heurísticas, 236, 238-239.
Estratificación, 35.
Evolución, 48, 58, 76, 81-88, 94, 229, 236-237.
Experimento de Ives-Stilwell, 167.
Experimento de Kennedy-Thorndike, 167.
Explicación estadística, 55-57, 58, 83, 85.
Explicación parcial, 57-59.
Explicación teleológica, 48, 81-82, 95.
Expresiones modales, 38, 127-128, 221.
Extensión, 29.

F

Farrel, B. A., 92-93, 95.
Feighl, Herbert, 22, 24, 72-73, 132-134, 140, 155.
Fenomenalismo, 106, 108, 118, 119 (nota), 121, 133.
Feyerabend, P. K., 25, 64-73, 113 (nota), 120-121, 125, 132-3, 157, 188, 247-248.
Fisher, R. A., 84 (nota), 85.
Fitch, V. L., 197.
Flew, Antony, 73, 141 (nota), 156.
Flogisto, 71.
Fodor, J. A., 91, 95, 248.
Fourier, J. B. J., 60.
Frecuencia, 40-44.
Funciones veritativas, 20, 203.

G

Galileo, 51, 69, 121.
Gardner, Martin, 193.
Gelernter, H. L., 237.
Gentzen, G., 240.
Geodésica, 189.
Geología, 15, 74.
Geometría, 183-193, 231-232, 237-238.
Geometría no euclidiana, 23, 106, 189-191, 193.
George, F. H., 247.
Gibbs, J. W., 62.
Gödel, Kurt, 29, 32-33, 34, 46, 116-117, 238, 240-241, 248.
Gold, T., 212, 213-214, 215-216, 217, 219.
Good, I. J., 46.
Goodman, Nelson, 53, 144, 146, 156, 248.

Griegos, 158.
Grünbaum, A., 25, 153-154, 157, 168, 181-182, 186-189, 192-193, 195, 206, 207 (nota), 218.
Guinea, Golfo de, 151.

H

Hacking, Ian, 145 (nota), 155.
Hamblin, C. C., 31 (nota).
Hamlyn, D. W., 248.
Hammond, L. M., 45.
Hanson, N. R., 24, 133.
Harlow, H. F., 95.
Harré, R., 133.
Harrod, R. F., 156.
Hayek, F. A., 95.
Hebb, D. O., 95.
Hedrick, E. R., 193.
Heisenberg, W., 19, 122, 124.
Hempel, C. G., 24, 46, 49, 62, 65, 71 (nota), 72, 118, 132, 157.
Henderson, L. J., 89, 95.
Hilbert, David, 45, 190-191, 234 (nota).
Hipótesis cuasi-ergódica, 210.
Hipótesis del continuo, 34.
Hipótesis efectiva, 148-149.
Hipótesis de Ptolomeo, 118, 121.
Historia natural, 74-79, 94.
Hobart, R. E., 247.
Hospers, John, 73.
Huellas, 196-197, 200, 213, 216-217, 218.
Hume, David, 14, 23, 110 (nota), 135-140, 141, 145, 156.

I

Idealismo, 21, 103.
Identidad, 31.
Indeterminismo, 173, 220-221, 223-224, 226-227.
Inducción, 105, 135-152, 155-156, 188, 237.
Inducción eliminativa, 138.
Información, 201-203, 218, 230, 231, 236, 241, 242, 245.
Instrumentalismo, 106, 108-109, 111-112, 117-118, 132-133, 150, 152.
Instrumentos calibrados, 102.
Inteligencia, 231-234, 236, 237-238, 244, 247-248.
Inteligencia artificial, 232-238, 247-248.
Interpretación de Copenhague, 22, 121-122, 123-124, 133.

J

Jakobson, R., 219.
Jammer, Max, 192.
Jeffress, L. A., 248.
Jeffrey, G. B., 192.
Jeffret, Richard C., 156.
Juegos, teoría de los, 16.
Júpiter, 55 (nota), 196.

K

Kant, Immanuel, 23, 158, 165, 246-247
Katz, Jerrold J., 141, 154, 155, 248.
Kelvin, 62.
Kemeny, 65, 156.
Keynes, J. M., 138, 155.
Kleene, S. C., 45.
Kneale, W., 46, 73, 155.
Körner, S., 25, 133.

L

Lammarck, J. B. de, 82.
Laplace, P. S. de, 215, 220-221, 223-224.
Leckie, G. G., 45.
Leibniz, Gottfried von, 159, 163-164, 182-183, 193.
Leverrier, U. J. J., 153.
Levi, I., 155.
Ley de Bode, 55.
Ley de Boyle, 59, 98-99.
Ley de Ohm, 96, 103.
Leyes de conservación, 124-125.
Ley de Snell, 97.
Leyes estadísticas, 58, 136, 148 (nota).
Leyes experimentales, 96-97.
Libre albedrío, 225, 226-230.
Línea del mundo, 170, 171-172, 176, 179, 189, 194.
Lingüística, 13, 17, 246-248.
Lobachevski, Nikolai, 189.
Lógica: de predicados, 28-31, 115; sentencial, 26-28, 115.
Lorentz, H. A., 167-168, 192.
Loschmidt, J., 199.
Lowinger, Armand, 157.

M

McCarthy, J., 247.
McCrea, W. H., 193.
Mach, Ernest, 23, 108, 110 (nota), 133, 162, 180, 182, 193, 213.

MacKay, D. M., 230.
McKinsey, J. C. C., 127.
Macroentropía, 204-205, 213.
Macro-fenomenalismo, 112, 119-120.
Magia, 22.
Magnitud derivada, 100-101.
Magnitud fundamental, 100-102.
Mainx, Félix, 95.
Marcus, Ruth Barcan, 46.
Margenau, H., 126 (nota), 133.
Marte, 55 (nota).
Maxwell, Grover, 25, 72-73, 119, 132-134, 157, 247.
Maxwell, James Clerk, 166, 167, 209.
Mayo, Bernard, 253.
Mecánica cuántica, 43, 57, 61-62, 109, 122-126, 131, 184, 197, 221, 223-224, 226-227, 229.
Mecánica estadística, 64-65, 197-201, 209, 210.
Mecanismo intencional, 91-92, 241-244, 247-248.
Medición, 96-101, 132.
Mehlberg, Henryk, 104, 132, 218.
Mellor, D. H., 132.
Memoria, 194-196, 199-200, 218, 232, 236-237.
Mercurio, 55 (nota), 68, 152, 154.
Metalenguaje, 36, 130-131, 241.
Metalógica, 39.
Metamatemática, 50, 130, 221, 240.
Michelson - Morley, experimento de, 167-168.
Miller, Dickinson, S., 247.
Minkowski, H., 168, 169, 171-172, 176, 178-179, 184, 192.
Minski, M. L., 237-238, 247.
Mischel, Theodore, 248.
Mises, R. von, 46.
Modelo, 118.
Modelo de explicación legal de cobertura, 57-59, 196.
Mónadas, 159.
Morgenbesser, S., 25, 133.
Movimiento browniano, 67.

N

Nagel, Ernest, 24, 46, 65, 71 (nota), 72, 73, 95, 132, 134, 155, 156, 191, 193.
Neidorf, Robert, 134.
Neptuno, 55 (nota).
Neumann, J. von, 33, 223, 247, 248.
Newell, Allen, 234, 235-236.
Newman, J. R., 46.
Newton, Isaac, 17, 48, 51, 60, 65-66, 68, 70-73, 85, 104-105, 127-128, 128 (nota), 137, 149, 153-154, 159-161, 162-165, 168-171, 182, 193, 215, 223.
Noble, C. A., 193.

O

Olbers, H. W. M., 62, 212-213, 217.
Ondas: avanzadas, 209; retardadas, 209-210.
Oparin, A. I., 95.
Operacionismo, 106, 108-110, 133, 163, 184.
Oppenheim, P., 65.
Optica geométrica, 65, 112.

P

Paleontología, 82.
Pap, A., 24, 134, 155.
Paradoja del reloj, 173-180.
Paridad, 165, 197.
Pascal, Blaise, 141 (nota).
Passmore, J. A., 156.
Pasado, presente y futuro, 194-195, 211.
Pauli, 126.
Péndulo de Foucault, 161-162.
Platón, 158.
Plutón, 55 (nota).
Política contra-inductiva, 142-143.
Popper, K. R., 24, 42-43, 46, 132, 133, 149-154, 156, 193, 210, 217, 218-219, 222.
Posibilidad, 40-41.
Pragmatismo, 49-50.
Predicados ostensivos, 144-145.
Presley, C. F., 133.
Príncipe, Golfo de Guinea, 151.
Probabilidad: interpretación de la frecuencia, 40-44, 46; interpretación de la tendencia, 42-45, 46, 124, 136; interpretación metalingüística, 39, 41-42.
Procedimiento de decisión, 30, 31, 46, 110 (nota), 146 (nota), 232, 233, 239.
Programa, 226-228, 231, 234-236, 237, 241, 243-244.
Progresión, 34.
Propiedad disposicional, 131-132.
Psicolingüística, 246, 247-248.
Psicología, 89-94, 246.
Pura casualidad, 224-225, 226.
Putnam, Hilary, 46, 65, 69-70, 133-134, 148, 149, 156, 188, 193, 240, 248.

Q

Quasar, 65.
Quine, W. V., 32, 35, 37 (nota), 45-46, 69, 127-128, 134, 154, 157, 192 (nota).

R

Ratoosh, P., 132.
Razones, 226-227, 230.
Razón suficiente, 163-164.
Redescripción, 59.
Registros (récords), 84, 196, 197, 205, 213, 216.
Registros indicadores (pointer readings), 103.
Regla *a priori*, 143.
Regla convergente, 142, 143.
Reglas de correspondencia, 64.
Regla directa, 142, 143, 146.
Regla oblicua, 142, 143.
Reglas de tipos, 35.
Reichenbach, 25, 73, 134, 140, 148, 155, 186, 192, 204, 206, 214, 218.
Relatividad, 62, 65, 68, 104 (nota), 105, 109-110, 128 (nota), 164-182, 213 (nota), 216.
Requisito epistemológico, 57.
Riemann, G. F. B., 182.
Roberts, K. V., 223.
Robertson, H. P., 193.
Rose, J. Holland, 249.
Rosser, Barkley, 33 (nota), 46, 240 (nota).
Routley, Richard, 128 (nota).
Russell, Bertrand, 35, 140, 155, 233, 234 (nota).
Ryle, Gilbert, 19-20, 44, 111, 132, 222 (nota).

S

Salmon, Wesley, 141, 143-147, 155.
Saturno, 55 (nota).
Scheffler, Israel, 16-18, 24.
Schilpp, P. A., 46, 156, 193.
Schlegel, Richard, 218.
Schlick, Moritz, 194.
Schröndinger, E., 197, 218.
Sciama, D. W., 223.
Scriven, Michael, 50, 58, 73, 85, 95, 103-105, 132, 134.
Sellars, Wilfrid, 66 (nota), 72, 108, 132-134, 248.
Semántica, 33, 35-36, 50, 202, 241, 245.
Sentencias condicionales contrarias a los hechos, 54.
Sentencias condicionales subjuntivas, 130, 221.
Sentencias de observación, 21, 107, 112-113, 117.
Sentencia de Ramsey, 113-115.
Shakespeare, William, 232.
Shannon, C. E., 201, 219, 247.
Shapere, Dudley, 72.
Shaw, J. C., 234, 235-236.
Simón, Herbert A., 234, 235-236, 247.
Sintaxis, 45.
Sistema derivado, 206, 207, 213-214, 216-217.
Sistema inerte, 128, 161-164, 180.
Skinner, B. F., 95, 237, 248.
Smart, J. J. C., 25, 72, 133, 134, 192, 247-248.
Smullyan, R., 45.
Sobral, Brasil, 151.
Sokal, Robert R., 77.
Spector, Marshall, 133.
Steinhardt, F., 45.
Stevens, S. S., 132.
Stove, D. C., 156.
Strawson, P. F., 156, 246 (nota).
Sugar, A. C., 127.
Suppes, P., 127.
Sintaxis, 32, 46, 50, 113, 130, 202, 221, 223, 241.

T

Tabla de verdad, 26-28.
Tarski, Alfred, 31, 35.
Taxonomía, 75-78.
Taxonomía estadística, 77 (nota).
Taylor, Richard, 242, 243, 248.
Teichrow, D., 245.
Temperatura, 69-70.
Términos estructurales, 16-18.
Termodinámica, 20, 64, 68, 104, 165, 197-201, 205, 213-219.
Teorema de Pitágoras, 190.
Teorema ergódico, 199.
Teorema H, 201, 218.
Teoría calórica, 147.
Teoría de Conjuntos, 33, 127, 238.
Teoría de la Medición, 186.
Teoría de la prueba, 36.
Teoría elemental de números, 32-33, 116, 238. 240.
Thomson, G. P., 72.
Thomson, Judith Jarvis, 156.

Thomson, W., 62.
Thornton, J. B., 133.
Tiempo, 158, 159-160.
Tiempo absoluto, 159-160, 179.
Tiempo relativo, 160, 192.
Tierra, 55 (nota), 77-78, 161.
Topología, 179, 186, 188-190, 212 (nota), 231.
Toulmin, Stephen, 24, 95, 111-112, 132, 199 (nota).
Traducción mediante máquinas, 245-246.
Transformación de Lorentz, 168-171.
Transformación galileana, 166-167, 169.
Turing, A. M., 46, 239-241, 244, 248.
Turlay, R., 197.
Turner, Merle B., 95.

U

Uranio, 153.
Urmson, J. O., 156.

V

Vacío, 158.
Valencia, 119.
Validación, 22, 140, 145, 150.
Van Kampen, N. G., 218.
Van der Waal, ecuación de, 98-99.
Variable: ligada, 28; libre, 28.
Variedad limitada, 138, 155.
Venus, 55 (nota), 161.
Verdad, 35-36.
Vindicación, 22, 140-141, 145, 148-149, 155.
Vulcano, 154.

W

Wald, George, 95.
Wallace, John R., 156.
Walter, W. Grey, 242.
Wartofsky, M. W., 46, 72, 248.
Watkins, J. W. N., 20-21, 156.
Watson, W. H., 18, 20-21, 24.
Weaver, Warren, 219.
Wells, Rulon, 219.
Weyl, H., 192.
Wheeler, J. A., 132, 164, 183-191.
Whitehead, A. N., 126 (nota), 233, 234 (nota).
Wiener, Norbert, 208, 218.
Wigner, Eugene, P., 193.
Will, F. L., 156.
Williams, D. C., 156.
Williams, W. T., 88, 95.
Wisdom, J. O., 156.
Wittgenstein, Ludwig, 17, 18-19, 25, 95, 140, 230.
Woodger, J. H., 95.
Woolsey, C. N., 95.
Wright, G. H. von, 46, 155, 156.

Z

Zenón, 191, 193.
Zermelo, 33.

OTRAS PUBLICACIONES DE EDITORIAL TECNOS

SERIE DE FILOSOFIA Y ENSAYO

Dirigida por
MANUEL GARRIDO

ABELLÁN, José Luis: *Miguel de Unamuno a la luz de la psicología.* Una interpretación de Unamuno, desde la psicología individual.

ABELLÁN, José Luis: *Ortega y Gasset en la filosofía española.* Ensayos de apreciación.

DÍAZ, Elías: *Revisión de Unamuno.* Análisis crítico de su pensamiento político.

DOPP, Joseph: *Nociones de lógica formal.*

ENJUTO BERNAL, Jorge: *La filosofía de Alfred North Whitehead.*

FERNÁNDEZ DE CASTILLEJO, José Luis: *Actualidad y participación.* Una filosofía contemporánea.

GARRIDO, M., y otros: *Filosofía y ciencia en el pensamiento español contemporáneo (1960-1970).*

GARRIDO, M., y otros: *La filosofía científica actual en Alemania.*

GARRIDO, M.: *Lógica simbólica.*

GURMÉNDEZ, Carlos: *Ser para no ser.* Ensayo de una dialéctica subjetiva.

HIERRO, José S.-P.: *Problemas del análisis del lenguaje moral.*

HUGHES, John B.: *José Cadalso y las «Cartas Marruecas».*

LAKATOS, Imre: *Historia de la ciencia y sus reconstrucciones racionales.*

LORENZEN, Paul: *Metamatemática.*

LORENZO, J. de: *La filosofía de la matemática de Jules Henri Poincaré.*

LORENZO, J. de: *Introducción al estilo matemático.*

MARTÍN SANTOS, Luis, y otros: *Ensayos de filosofía de la ciencia.* Simposio de Burgos. En torno a la obra de Karl R. Popper.

MATES, Benson: *Lógica matemática elemental.*

MOULOUD, Noël: *Lenguaje y estructuras.* Ensayos de Lógica y Semiología.

NICOL, Eduardo: *El problema de la filosofía hispánica.*

NICOL, Eduardo: *Historicismo y existencialismo* (2.ª ed.).

PARÍS, Carlos: *Hombre y naturaleza.*

QUINE, W. V.: *La relatividad ontológica y otros ensayos.*

QUINTANILLA, Miguel A.: *Idealismo y filosofía de la ciencia.* Introducción a la Epistemología de Karl R. Popper.

RAMA, Carlos M.: *Teoría de la historia.* Introducción a los estudios históricos (3.ª ed.).

RODRÍGUEZ PANIAGUA, José M.ª: *Hacia una concepción amplia del Derecho natural.*

RODRÍGUEZ PANIAGUA, José M.ª: *Marx y el problema de la ideología.*

SÁNCHEZ DE LA TORRE, Angel: *Los griegos y el Derecho natural.*

SOTELO, Ignacio: *Sartre y la razón dialéctica.*

TIERNO GALVÁN, Enrique: *Acotaciones a la historia de la cultura Occidental en la Edad Moderna.*

TIERNO Y GALVÁN, Enrique: *Razón mecánica y razón dialéctica.*

COLECCION VENTANA ABIERTA

AGUILAR NAVARRO, M.; ALONSO GARCÍA, M.; BERNÁRDEZ CANTÓN, A.; CARRILLO, J. A.; DÍEZ DE VELASCO, M.; ECHEVARRÍA, J.; GIMÉNEZ FERNÁNDEZ, M.; JIMÉNEZ DE PARGA, M.; LISSARRAGUE, S.; LUCAS VERDÚ, P.; MARAVALL, J. A.; TRUYOL, A.: *Comentarios universitarios a la «Pacem in Terris»*.

BELTRÁN, Lucas: *La integración económica europea y la posición de España.*

BENEYTO, Juan: *La opinión pública.* Teoría y Técnica (2.ª ed.).

BENEYTO, Juan: *La opinión pública internacional.*

ECHEVARRÍA, Juan: *Anotaciones al Plan de Desarrollo.*

GONZÁLEZ CASANOVA, J. A.: *Elecciones en Barcelona* (1931-1936).

JENKS, C. Wilfred: *La justicia social en el Derecho de las naciones.* El impacto de la O. I. T. después de cincuenta años.

JIMÉNEZ DE PARGA, Manuel: *Formas constitucionales y fuerzas políticas.*

JIMÉNEZ DE PARGA, Manuel: *La V República Francesa.*

JIMÉNEZ DE PARGA, Manuel: *Las monarquías europeas en el horizonte español.*

LUCAS VERDÚ, Pablo: *Política e «Inteligencia»* (2.ª ed.).

MARTÍN-RETORTILLO, Lorenzo: *Libertad religiosa y orden público.*

MCNAMARA, R.: *Cien países, dos mil millones de seres.* La dimensión del desarrollo.

NOEL-BAKER, Philip: *La carrera de armamentos.*

PELLING, Henry: *El sindicalismo norteamericano.*

RAMA, Carlos M.: *La Historia y la novela* (2.ª ed.).

ROSTAND, Jean: *Ciencia falsa y pseudociencia.*

THORP, Willard: *La literatura norteamericana en el siglo XX.*

TRUYOL Y SERRA, Antonio: *Los derechos humanos.*

YOUNG, Michael: *El triunfo de la meritocracia.*

GINER DE LOS RÍOS, F.; AZCÁRATE, G. de; SALMERÓN, N.: *La cuestión universitaria. Epistolario.* (Estudio preliminar de Pablo de Azcárate.)

EXISTEN, dice el autor de este libro, tres tipos de textos o tratados sobre la Filosofía de la Ciencia. En primer lugar, está la Filosofía de la Ciencia del filósofo; en segundo lugar, la Filosofía de la Ciencia, que se dirige a los estudiantes de Filosofía que no tengan formación científica alguna; y en tercer lugar, está la Filosofía de la Ciencia del científico, que se dirige a quienes tienen cierta preparación científica y suficiente preparación filosófica. «Entre Ciencia y Filosofía» puede clasificarse en este último grupo. Es una introducción a la Filosofía de la Ciencia en la que se discute tanto la Ciencia y su Metodología como los resultados de la Ciencia. Escrito el libro desde el punto de vista científico, el profesor Smart se refiere constantemente a los problemas actuales de la Ciencia, estudiando sus relaciones con los problemas que preocupan a los filósofos, en particular aquellos que atañen a los dos campos, por tratar de cuestiones interdisciplinarias, en las que Filosofía y Ciencia intercambian conceptos y métodos.

Las principales cuestiones que se discuten en el libro que ofrecemos a nuestros lectores, se refieren al contenido general de la Filosofía de la Ciencia, a la legalidad en las distintas campos científicos y al método de la Ciencia, con especial atención respecto de la tesis de Grünbaum acerca del Espacio y del Tiempo.

Dentro de su relativa sencillez, es un libro inexcusable para estar al día en cuanto se refiere a los problemas actuales de la Filosofía y de la Ciencia.